University of Alberta

Functional Analysis of Deletion Mutants of FinO, the Fertility Inhibition Protein of the F-like Plasmids

by

James R. Sandercock

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE.

DEPARTMENT OF BIOLOGICAL SCIENCES

Edmonton, Alberta
Spring, 1998

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-28985-0

Canada

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Functional Analysis of Deletion Mutants of FinO, the Fertility Inhibition Protein of the F-like Plasmids** submitted by James R. Sandercock in partial fulfillment of the requirements for the degree of Master of Science.

Dr. Laura Frost, Supervisor

Dr. John Bell

Dr. George Owttrim

Dr. Mark Glover

Date: Nov. 5, 1997

To my parents, Bob and Libby, who's love and support are unending.

~

To my sister, Knansie, who (incessantly) challenges my world-view.

Abstract

High frequency conjugative transfer of the F-like plasmids is possible when TraJ protein, the main activator of transcription of the transfer operon, is synthesized in the donor organism. Transfer is negatively regulated by the FinP antisense RNA, which hybridizes to the positive sense transcript, TraJ, thereby blocking TraJ protein synthesis. The FinO protein inhibits transfer of the plasmid by protecting FinP RNA from degradation, and by promoting duplex formation between the FinP and TraJ transcripts. The purpose of this thesis is to delineate the minimal sequence required for fertility inhibition, as mediated by FinO, and to define the regions of the protein that are responsible for the above functions.

The *finO* gene and downstream sequences from plasmid R6-5 were cloned into the multicopy plasmid pBC-KS+. A series of nested deletions of the downstream sequences were created, and the deletion products were subcloned. Clones which contained the deletion products were assayed by a number of means. The results indicate that the downstream sequences do not play a role in fertility inhibition, and that minor deletions of the 3' end of the *finO* gene result in the loss of fertility inhibition.

Fragments of the *finO* gene were translationally fused to the 3' end of the *gst* gene of the expression vector pGEX-2T. The *gst*-fusion constructs were examined for their ability to repress conjugal transfer, and for their ability to stabilize FinP RNA *in vivo*. GST-fusion peptides were purified and assessed for their ability to bind FinP RNA and promote duplex formation between FinP and TraJ transcripts *in vitro*. The specificity of binding was also examined.

GST-FinO protein bound FinP RNA with high affinity, with an association constant (k_a) of 5.6×10^7 M^{-1}. Binding specificity assays indicated that GST-FinO bound the target RNA, FinP, with a moderate degree of specificity. The protein increased the rate of duplex formation between the transcripts ~3.7-fold *in vitro*. The plasmid which encodes GST-FinO, pGEX-FO2, reduced mating efficiency ~300-fold, and stabilized FinP *in vivo*, with a half life > 60 minutes. The fusion product GST-FinO73, which includes the N-

terminal 73 amino acids of FinO, bound FinP with a k_a of ~2.2 x 10^7 M^{-1}, and displayed a moderate specificity for the FinP target. GST-FinO73 promoted duplex formation ~3-fold *in vitro*. The peptide conferred no stability to FinP transcripts when expressed *in vivo* from the plasmid pGEX-FO73. The intermediate sized peptide, GST-FinO141 bound FinP with a k_a of 3.5 x 10^7 M^{-1}, and displayed a moderate specificity for the target RNA. The peptide increased the duplex formation rate 19-fold *in vitro*. The peptide did not confer fertility inhibition or FinP stability when expressed from the plasmid pGEX-FO141 *in vivo*.

From these data a model was developed which proposes a molecular mechanism by which FinO protein protects FinP from degradation, and by which FinO promotes duplex formation between FinP and TraJ, thereby repressing expression of the transfer operon.

Acknowledgements

First and foremost I want to thank my advisor, Dr. Laura Frost, for giving me a place in the lab to hang my hat. Throughout the first few frustrating years you have been much more than supportive. Thank you for talking me down from the ledge on more than one occasion. Oh, and thanks for the Scotch. Thanks also go out to Tim van Biesen, who preceded me on the FinO project. Not only did he blaze the GST-FinO trail, but he also left several constructs, which proved to be invaluable for much of my research.

Karen Anthony has been of great help throughout my stay in the lab. Karen, you make an excellent senior graduate student, and you've contributed to my molecular tutelage in many ways. Lori Jerome has been my partner in crime on the FinOP system. She introduced me to the big bad world of RNA. By my estimation I'd have accomplished very little without her laboratory expertise and overall level-headedness. Lori, there is not enough thanks by which I can give you proper credit. Big Bill Klimke has been the "Frost Lab Guru" of wit, sarcasm, and computers for the last three years. Thanks Bill for your help with GCG, internet stuff, and Civ. I also extend thanks to Jan Manchak, Richard Fekete, John Simon and the many summer students/special project students who made my stay in the Frost Lab enjoyable.

To my science mentors, Tony Cornish, Brent Rudy and Linda. Thank you for showing me the ropes in my undergraduate years. You taught me much about science, and it's proper place in one's life. Kudos to Dr. Roy, who (unceremoniously) kicked my hind end into high gear in my third year of undergraduate studies. Dr. Roy, Tom Hantos and Tony Cornish, you have shown me through word and deed how to be a professional and excellent instructor.

Finally, and with considerable fan-fare... thank you to the Young Life community. Colin, Gilles, Anthony and Kevin, you are the ones who kept me sane throughout my years of graduate studies. Thank you Gilles for adopting me into the fold! Thanks also to the YL "kids" who's friendships have been some of the most meaningful. And Steve... the thesis would have never flown without your help with Adobe.

Table of Contents

	Page
Chapter 1. Introduction	
A. Conjugation	2
I. Classification of plasmids	5
II. The F plasmid	6
B. Regulation of Transfer in F	11
I. TraJ, the main activator of the *tra* operon	12
II. FinP, an antisense RNA	15
III. FinO, an RNA binding protein	20
IV. Interaction of FinO with FisO RNA	24
C. Regulation of Transfer in R100 and R100-1	24
I. R100 FinP and TraJ RNA	25
II. TraM and FinM transcripts as anti-antisense RNAs	28
III. The role of FinO in R100	29
IIIa. FinO as an antisense RNA	29
IIIb. FinO protein as a regulator of *traM* expression	31
D. Other Sense-Antisense RNA Systems	32
I. Replication control of plasmid ColE1	33
II. Copy number control in IncFII plasmids	37
III. Copy number control in pT181 and pIP501	40
IV. Control of transposition events of the Tn*10*/IS*10* elements	43
E. RNA binding proteins	44
I. Rom/Rop protein of plasmid ColE1	45
II. RNA binding and duplex promotion by the protein p53	46
III. Tat protein and TAR RNA of HIV-1 and HIV-2	47
IV. Rev protein and RRE RNA of HIV-1 and HIV-2	54
F. Summary	58
G. Research Objectives	59

Chapter 2. Materials and Methods

A. Strains and plasmids 62

B. Chemicals and enzymes 64

C. Exonuclease III deletions 64

D. Mating assays 65

E. β-galactosidase assays 66

F. Preparation of RNA and northern transfers 66

G. Construction of pGEX-fusion plasmids 68

H. GST-fusion protein purification 69

I. *In vitro* transcriptions 70

J. Mobility shift assays: RNA binding, competition assays 71

K. Duplex formation assays 71

L. Calculations 71

I. Calculation of the association equilibrium constant, k_a 71

II. Calculation of the binding rate constant, k_{app} 72

Chapter 3. Deletion Mutations of the Fertility Inhibition Protein FinO Cause Derepression of Transfer of the F plasmid

A. Introduction 75

B. Results 76

I. Exonuclease digestion of *finO* region of R6.5 76

II. Qualitative assays of FinO activity *in vivo* 81

III. Mating inhibition by R6.5 *finO* deletion derivatives 83

C. Discussion 84

Chapter 4. Characterization of the Functional Domains of FinO, the Fertility Inhibition Protein of F-like Plasmids.

A. Introduction 87

B. Results 87

I. Creation of truncated GST-FinO constructs 87

II. Conjugation is not inhibited by truncated GST-FinO products 91
III. GST-FinO derivatives do not stabilize FinP RNA *in vivo* 95
IV. The N-terminal α-helical domain is required for FinP binding 98
V. FinP RNA binding by GST-FinO derivatives is specific 110
VI. Duplex formation rate is improved by GST-FinO derivatives 115
C. Discussion 126

Chapter 5. Discussion
A. The role of downstream sequences in fertility inhibition 134
B. RNA binding activity of FinO 135
C. RNA protection activity of FinO 136
D. Promotion of duplex formation by FinO 137
E. A model of the functional domains of FinO 139
F. Direction of further studies 143

Chapter 6. References 145

List of Tables

Table	Title	Page
2.1	List of *Escherichia coli* K-12 strains employed	63
2.2	List of plasmids employed	63
3.1	Genotype and phenotypes of deletion derivatives of plasmid pBC 4.0	82
4.1	Mating efficiencies of plasmid pOX38-Km with co-resident plasmids	94
4.2	Summary of the functional properties of GST-fusions	125

List of Figures

Figure	Title	Page
1.1	Schematic diagram of plasmid mediated conjugation of gram negative bacteria	4
1.2	Schematic diagram of the F plasmid	8
1.3	Gene organization of the *tra* region	10
1.4	The regulatory region of the F *tra* region	14
1.5	The 2^{o} structure of FinP and TraJ RNA species	18
1.6	Alleles of *finO*	23
1.7	Organization of the *tra* regulatory region of R100 and R100-1	27
1.8.i	Regulation of primer formation at the ColE1 origin of replication	35
1.8.ii	Molecular basis for the hybridization of RNA I and RNA II	35
1.9.i	Regulation of the synthesis of replication factor RepA of plasmid R1	39
1.9.ii	Molecular basis for the hybridization of CopT and CopA RNAs	39
1.10	Molecular basis for the regulation of the synthesis of replication factor RepR of plasmid pIP501	42
1.11	Regulation of the synthesis of *IS*10 transposase	42
1.12	Schematic diagram of the TAR domain from HIV-1 and HIV-2	49
1.13	Diagram of the domains of RNA binding protein Tat	49
1.14	HIV TAR RNA structure when bound or unbound by arginine residues	53
1.15	Secondary structure of the RRE RNA of HIV	56
3.1	Diagram of map positions of deletion mutants of plasmid pBC 4.0, as produced by exonuclease III	78
3.2	Restriction map of the exonuclease III-generated mutants of the plasmid pBC 4.0	80

4.1	The predicted secondary structure of the FinO protein	89
4.2	10% SDS-polyacrylamide gel of GST-fusion proteins prepared from *E. coli* DH5α	93
4.3	Northern hybridization analysis: pGEX-FO2, -2T, -FO141	97
4.4	Northern hybridization analysis: pGEX-FOC37, -FO73, -FO34	100
4.5	Association of GST-FinO with [^{32}P]-labeled FinP or TraJ184 RNA	103
4.6	Association of GST-FinO141 with [^{32}P]-labeled FinP or TraJ184 RNA	103
4.7	Association of GST-FinO73 with [^{32}P]-labeled FinP or TraJ184 RNA	105
4.8	Co-incubation of GST-FinO34 with [^{32}P]-labeled FinP RNA	107
4.9	Co-incubation of GST-FinOC37 with [^{32}P]-labeled FinP RNA	107
4.10	Co-incubation of glutathione S-transferase with [^{32}P]-labeled FinP RNA	109
4.11	Specificity of binding by the protein GST-FinO	112
4.12	Specificity of binding by the protein GST-FinO73	114
4.13	Specificity of binding by the protein GST-FinO141	117
4.14	Duplex formation between [^{32}P]-labeled FinP and [^{3}H]-labeled $TraJ_{184}$ transcripts (GST and blank)	120
4.15	Duplex formation between [^{32}P]-labeled FinP and [^{3}H]-labeled $TraJ_{184}$ transcripts (GST-FinO)	122
4.16	Duplex formation between [^{32}P]-labeled FinP and [^{3}H]-labeled $TraJ_{184}$ transcripts (GST-FinO73 and GST-FinO141)	124
5.1	A schematic diagram of the proposed model for the activities of the domains of protein FinO	141

Abbreviations

A_{420}	absorbance at 420nm
ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleic triphosphates
EDTA	ethylenediaminetetraacetic acid
F	F conjugative plasmid
GST	glutathione S-transferase
IPTG	isopropyl β-D-thiogalactopyranoside
IS	insertion sequence
kDa	kilo-Daltons
M	molar concentration
OD_{600}	optical density at 600nm
ONPG	o-nitrophenyl β-D-galactopyranoside
orf	open reading frame
PAG	polyacrylamide gel
PCR	polymerase chain reaction
PEG	polyethylene glycol
RNA	ribonucleic acid
mRNA	messenger RNA
tRNA	transfer RNA
RNase	ribonuclease
SDS	sodium dodecyl sulphate
sec; s	seconds
Tn	transposon
UTP	uridine triphosphate

Chapter 1. Introduction

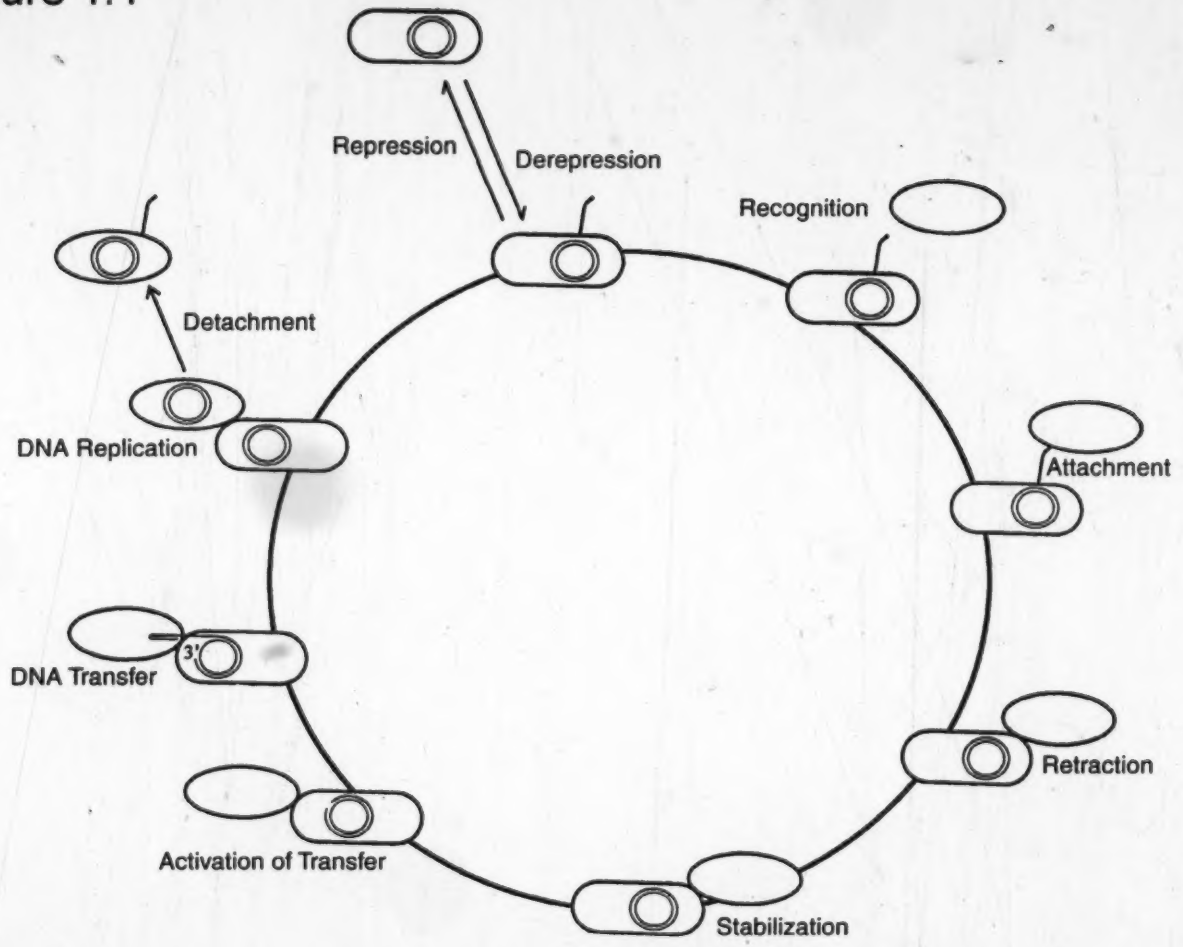
A. Conjugation

The horizontal transmission of genetic information between bacteria was first reported by Lederberg and Tatum (1946), who considered the process a bacterial sexual event. Initially called the Fertility, or F factor, the element responsible allowed for efficient transfer and integration of chromosomal markers from one bacterium to another. The element was later determined to be a circular extrachromosomal DNA of 100 kb size (Marmur *et al.*, 1961; Freifelder, 1968; Johnson and Willetts, 1980). Many plasmids, which are closely related to F, were found in the years following the initial discovery. To date plasmids have been described in nearly every genus of eubacteria, many archaebacteria, and some eukaryotes. Plasmids are one of the most efficient natural mechanisms for the transfer of antibiotic resistance genes between species, genera, and even kingdoms of organisms. Many plasmids also carry virulence markers, potentially broadening the pathogenic range of their bacterial hosts.

Because conjugative plasmids are of such importance from a medical perspective, it is critical to determine the molecular basis by which plasmids regulate transfer. The specific events leading to plasmid transfer are well characterized in many systems, and many reviews exist (Willetts and Wilkins, 1984; Chapman and Carlton, 1985; Ippen-Ihler and Maneewannakul, 1991; Farrand, 1993). The most thoroughly characterized mechanisms of bacterial conjugation are those of the gram negative bacteria. A schematic representation of the stages of conjugal transfer of gram negative plasmids is shown in Figure 1.1. Long filaments called pili are synthesized and extruded from the host cell, extending from the host cell surface into the medium. The pili can adhere to recipient cells *via* interactions between the pilus tip and specific receptors on the recipient cell (Frost *et al.*,1994). Following contact, an unknown signal causes pilus retraction, which brings the cells into close proximity. There is some evidence that, if retraction is blocked, a single-stranded copy of the plasmid DNA can be transferred through the lumen of the unretracted pilus at very low frequencies (Harrington and

Figure 1.1 Schematic diagram of plasmid mediated conjugation of Gram negative bacteria. The stages of conjugative transfer and regulation of transfer are shown.

Figure 1.1


Repression
Derepression
Recognition
Attachment
Retraction
Stabilization
Activation of Transfer
DNA Transfer
3'
DNA Replication
Detachment

Rogerson, 1990). However, if the pilus is allowed to retract, the donor and the recipient cells are brought into immediate contact. Membrane-bound adherence proteins then stabilize the mating bacteria, and this results in a higher frequency of mating (Miki *et al.*, 1978; Firth and Skurray, 1992). An unknown signal triggers nicking of the plasmid at a specific site, termed the origin of transfer (*oriT* or *nic*). A single strand of plasmid DNA is then transferred in the 5' to 3' direction from the donor to the recipient (Wilkins and Lanka, 1993). Once the recipient has synthesized the complementary strand of DNA, it is considered a donor organism for the plasmid. Following conversion, donor cells produce plasmid-specific exclusion factors, which block pilus adherence to the cell surface of the organism (Sukupolvi and O'Connor, 1990; Achtman *et al.*,1980), decreasing the chances of "homosexual" transfer between donor cells. In the case of F, an additional exclusion factor, TraS, is synthesized. TraS acts by blocking the redundant transfer of DNA between donor organisms (Achtman *et al.*, 1977).

I. Classification of Plasmids

Plasmids were originally classified on the basis of the antigenicity of their pili. Current classification is based instead on incompatibility groupings (Scott, 1984). Closely related plasmids cannot co-exist within the same host because the plasmid replication systems compete for essential replication factors. As a result of this competition, one plasmid will replicate at a higher rate and/or at the expense of the other, eventually resulting in a bacterial culture which contains only one plasmid type (Novick and Hoppensteadt, 1978). Plasmids that have different mechanisms of replication will not compete for replication components, and so are considered compatible. These plasmids can be stably maintained as co-residents of the same organism.

The F plasmid and closely related plasmids belong to the incompatibility group IncF. The IncF group is subdivided into groups I (F, ColV, R453), II (ColB2, R1, R6-5, R100), III (pSU306), IV (R124), V (pED208), VI and VII (Ippen-Ihler and Skurray, 1993). Plasmids classified within the same subdivisions are

incompatible while plasmids from different IncF subdivisions are compatible. i.e. F and ColV are incompatible while F and R100 are compatible. While vegetative replication events can differ considerably between the IncF plasmids, their conjugative and regulatory components share considerable similarities. As a result the conjugative and regulatory events of F are frequently used as the model for all IncF plasmids.

II. The F plasmid.

The F plasmid (Figure 1.2) is organized to some degree upon domains of similar function. The leading region is the portion of the plasmid that is transferred into the recipient cell first (Ray and Skurray, 1983). The leading region is located next to the origin of transfer. Many of the maintenance genes are grouped in the leading region so as to be amongst the first genes expressed in the new bacterial host. The Rep regions encode the genes required for autonomous vegetative replication by the plasmid. RepFIA contains two origins of replication: *ori*V, a bidirectional replication origin, and *ori*S, a unidirectional replication origin (Lane, 1981). RepFIB is also a functional replicon of F (Firth *et al.*, 1996). Several transposons and insertion sequences are also found in F, and these profoundly influence the biology of the plasmid. Tn*1000*, which interrupts the RepFIC region (Saadi *et al.*, 1987) allows F to integrate into the host chromosome, giving rise to Hfr strains. IS*3* interrupts the *finO* gene, which is involved in fertility inhibition (Cheah and Skurray, 1986). As a result of IS*3*, the F plasmid is derepressed for transfer. Finally the region responsible for transfer of the plasmid is organized in what is commonly referred to as the *tra* region (see also Figure 1.3). The 35 kb *tra* region encodes nearly 40 genes, most of which are organized in a single operon behind the pY promoter, which is named for the first gene of the operon, *traY* (Mullineaux and Willetts, 1984; Maneewannakul *et al.*,1992). Three genes are located upstream of the pY promoter. These genes are important for the regulation of transfer, and each is under the control of its own promoter. *traM* encodes a 14.5 kDa protein which plays an essential role in relaxosome

Figure 1.2 Schematic diagram of the F plasmid, adapted from Firth *et al* (1996). Map units are indicated in kilobase (kb) pairs. Transfer of a single-stranded copy of F occurs from *oriT*, indicated by a small arrowhead, and proceeds in a counter-clockwise direction. Replication is mediated by RepFIA or RepFIB. Transfer is regulated and mediated by the products of the genes located within the transfer (*tra*) region. Several transposons and insertion sequences have been mapped to F. These are indicated as black boxes.

Figure 1.2

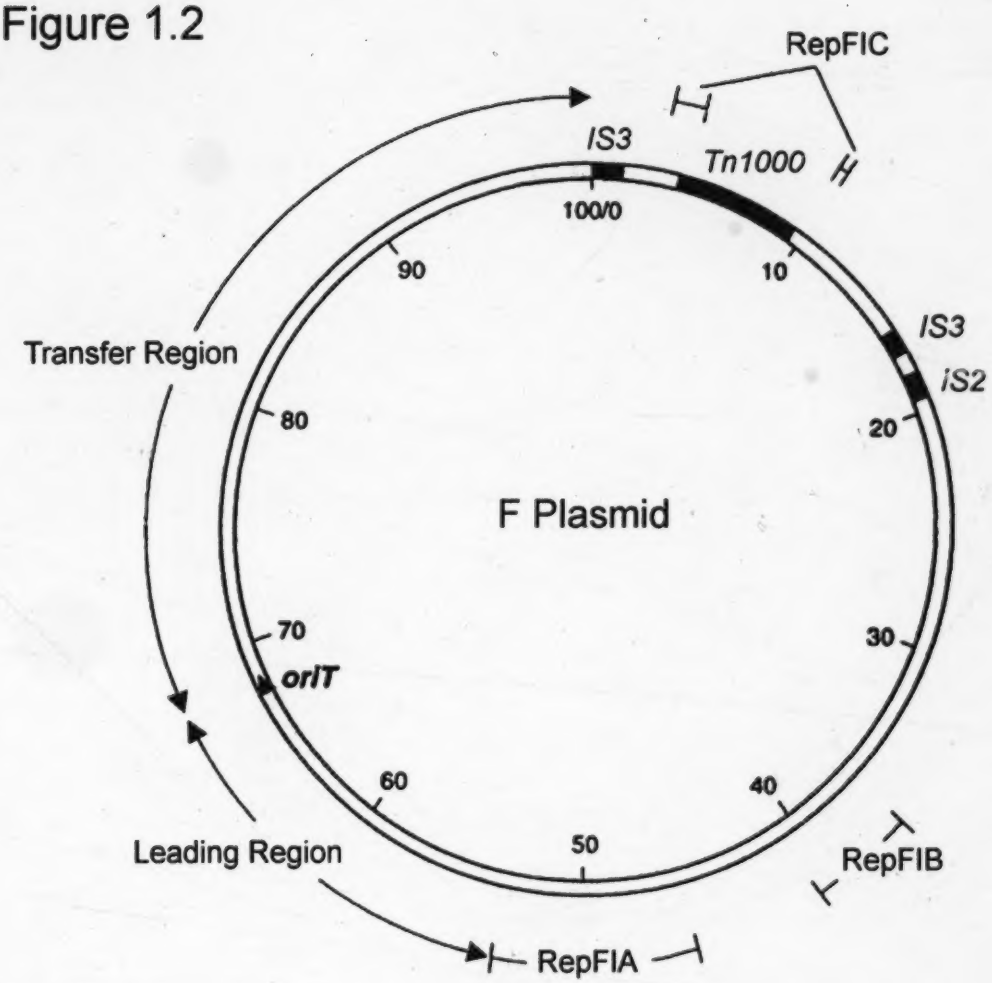
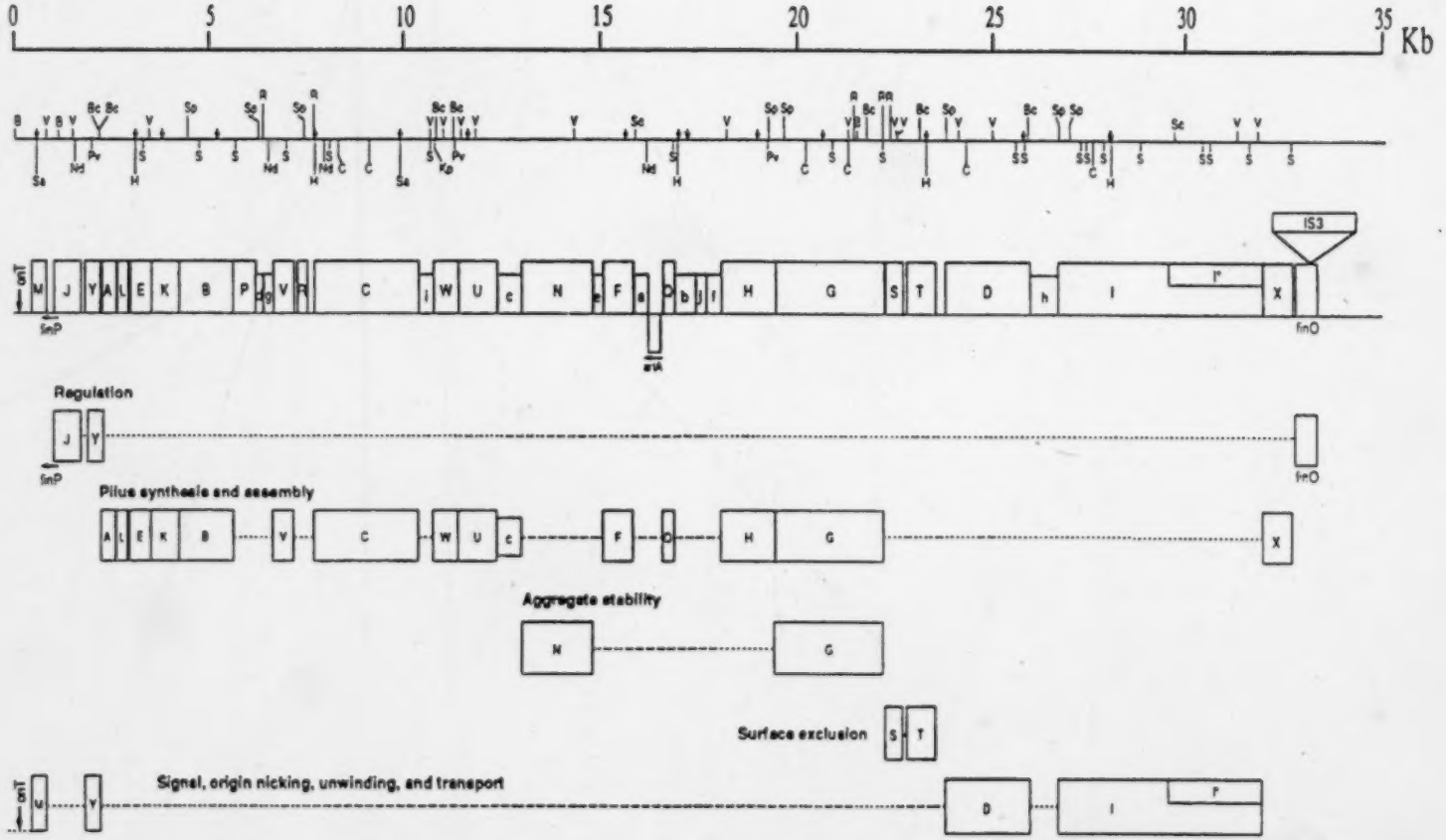

RepFIC
IS3
Tn1000
100/0
10
IS3
IS2
20
30
40
RepFIB
50
60
70
oriT
80
90
Transfer Region
Leading Region
RepFIA
F Plasmid

Figure 1.3 Gene organization of the *tra* region (Frost *et al.*, 1994). The top line indicates the length of the region in kb. 38 genes have been identified within the 35 kb region, and these are indicated on the third line. The majority of the genes are organized within a single operon, which is continuous from *traA* through to *traI* and perhaps beyond. *artA* and *finP* are both transcribed from the opposite strand. The *tra* genes have been classified by their functional roles in conjugal transfer, and are schematically grouped on the basis of these functions (lines four through eight).

Figure 1.3


0
5
10
15
20
25
30
35
Kb
IS3
finP
finO
traA
Regulation
Pilus synthesis and assembly
Aggregate stability
Surface exclusion
Signal, origin nicking, unwinding, and transport

formation at *oriT* (Di Laurenzio *et al.*,1992; Penfold, 1995). *traJ* encodes a 27 kDa protein which positively regulates transcription from pY (Silverman *et al.*,1991). *finP* is a gene which is transcribed from within the *traJ* leader region and is in the opposite orientation relative to the *traJ* gene (Finnegan and Willetts, 1971; Mullineaux and Willetts, 1985) . As a result, the FinP RNA product is antisense to the TraJ transcript. Finally, the *finO* gene is located at the distal end of the *tra* region. FinO protein acts by binding and protecting FinP RNA from degradation (van Biesen and Frost, 1994). It has been proposed that FinO protein may act with FinP RNA to block translation of the TraJ protein from available TraJ mRNA (van Biesen and Frost, 1994). This thesis will focus on the *tra* region and the molecular basis for the regulation of the *tra* operon by these gene products.

B. Regulation of Transfer in F

The regulation of transfer in F-like plasmids is based on a two component fertility inhibition system designated the FinOP system. For the most part the FinOP system is identical in all the F-like plasmids, although some very subtle regulatory differences do exist. The two components, an antisense RNA named FinP and an RNA binding protein, FinO, are believed to inhibit transfer by blocking the translation of the main activator of pY transcription, TraJ (van Biesen and Frost, 1994). In F, the fertility inhibition system is deregulated due to an interruption of the *finO* gene by an IS*3* element (Cheah and Skurray, 1986). As a result, F expresses the *tra* operon constitutively, and so transfers the plasmid in a derepressed, or uncontrolled manner. This lack of control can have dire consequences for the host, first because the constitutive production of transfer machinery (such as the pilus) is energetically expensive (Levin and Lenski, 1983) and second, because the permanent extracellular display of pili makes the host susceptible to pilus specific bacteriophage (Anderson, 1968). However, transitory derepression of transfer is observed with many plasmids (Lundquist and Levin,

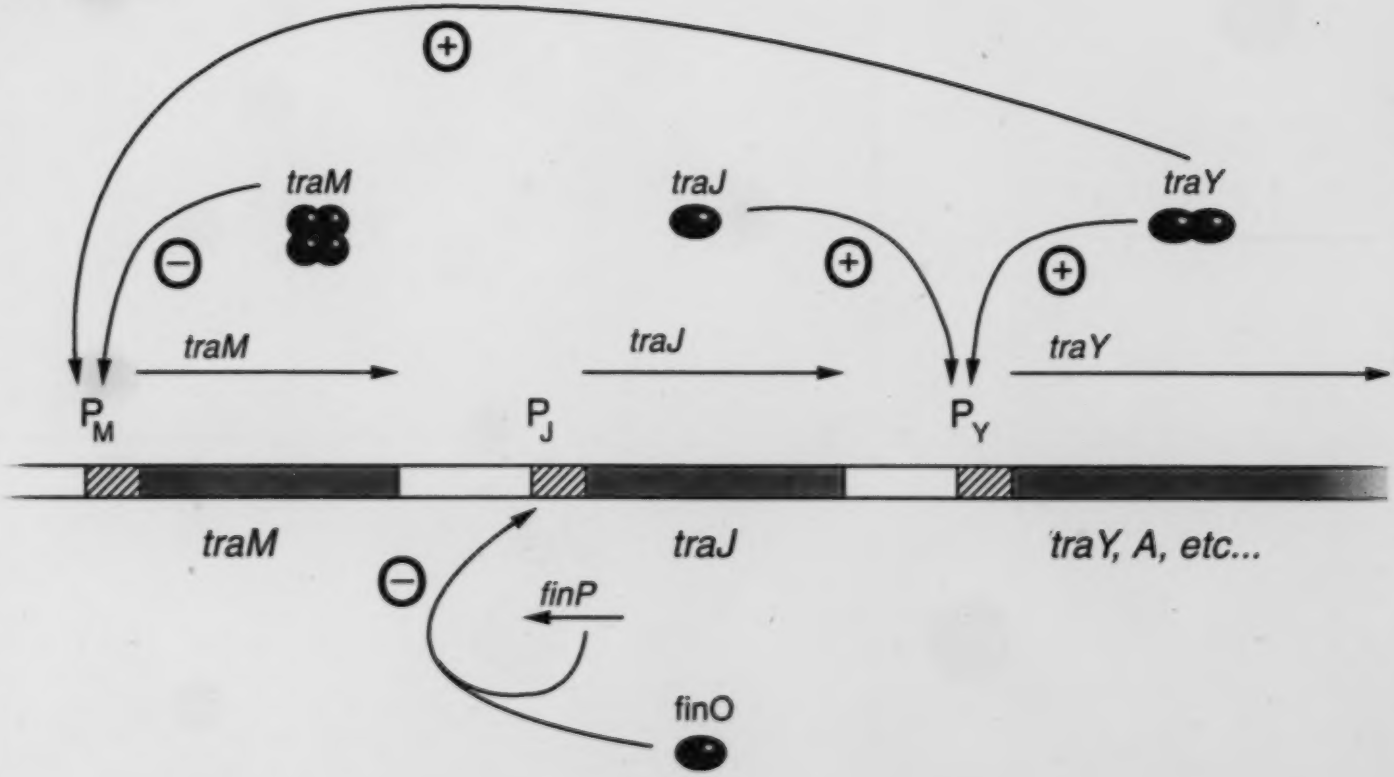
1986; Simonsen, 1990). The phenomenon of transitory derepression, also known as HFT for High Frequency of Transfer, is of particular importance for the rapid dissemination of the plasmid through a culture which is growing on solid media. Directly following transfer of the plasmid, the new recipient is derepressed for transfer, and rapidly infects neighboring cells with the plasmid. After several generations regulation of transfer is imposed. This system allows a plasmid to sweep through a bacterial population in epidemic proportions, but also avoid the negative results of permanent derepression of transfer (Simonsen, 1990). A general schematic for the regulation of transfer in IncF plasmids is shown in Figure 1.4 and is described in the text below.

I. TraJ, the Main Activator of the *tra* Operon

The main positive regulator of transfer by the F plasmid is the TraJ protein (Willetts, 1977). The TraJ protein is a 27 kDa cytoplasmic protein (Cuozzo and Silverman, 1986) which acts in concert with the SfrA, IHF and TraY proteins (Silverman *et al.*,1991; Nelson *et al.*,1993) to promote transcription from the pY promoter (pY). To date TraJ protein has not been observed to bind DNA directly, leading Silverman and colleagues to propose that pY is regulated by a complex of proteins, in which TraJ is the major activator (Gaudin and Silverman, 1993). This complex may cause a change in the superhelical density of the pY region, or may act more directly to facilitate RNA polymerase loading to the promoter region. Alternatively TraJ may be a sigma factor required for transcription initiation, though the protein shares little resemblance with known sigma factors. Currently there are five known alleles of TraJ (Willetts and Maule, 1986; Di Laurenzio *et al.*, 1991; Gras-Goldner *et al.*,1990). The only regions of significant homology between the *traJ* products are the untranslated leader region of the TraJ mRNAs, and a DNA binding helix-turn-helix motif at the protein level. The helix-turn-helix motif is shared by many DNA binding proteins, including sigma factors. Some pY activity has been observed in the absence of *traJ*, and so it is possible that other weak promoter(s) for the *tra* operon exist (Fowler and Thompson, 1986; Silverman *et al.*,1991; Carter and Porter, 1991).

Figure 1.4 The regulatory region of the F *tra* region. TraJ protein is synthesized from the *traJ* transcript. TraJ positively regulates transcription of the *traY* operon from the promoter, P_Y. TraY in turn positively regulates the P_Y and P_M promoters, and acts at the *oriT* with TraI to promote single-stranded nicking. TraM protein negatively regulates its own expression. Translation of TraJ protein is negatively regulated by *finP* RNA, which is labile *in vivo*. FinO protein stabilizes *finP* RNA, resulting in the repression of TraJ translation, and therefore represses expression from the *traY* operon.

Figure 1.4


traM
traJ
traY
traM
traJ
traY
P_M
P_J
P_Y
traM
traJ
traY, A, etc...
finP
finO

In addition to negative regulation of TraJ expression by the FinOP system (discussed below), the Cpx two-component regulatory system also negatively regulates TraJ expression (Silverman *et al.*, 1993). Recent work on the Cpx (Conjugative pilus expression) cascade indicates that the inner membrane sensor protein, CpxA responds to membrane stresses such as pH changes and the overexpression of the outer membrane lipoprotein NlpE (Pogiano *et al.*, 1997), the lack/loss of phosphatidylethanolamine (Mileykovskaya and Dowhan, 1997), and protein-mediated toxicity (Cosma *et al.*, 1995). The phosphorylated form of the response regulator protein, CpxR, is able to directly bind a putative Cpx consensus sequence (5' $GTAAN_{(6-7)}GTAA$ 3') and promote the expression of genes such as *degP* and *ppiA* (Pogliano *et al.*, 1997). It is uncertain at this time if CpxR directly regulates TraJ expression, or if the regulation of TraJ expression is part of a larger response to extracellular damage.

II. FinP, an Antisense RNA

FinP RNA is transcribed beginning at the second codon of the *traJ* coding region from the strand opposite the *traJ* gene. Transcription extends into the 105 nucleotide (nt) untranslated leader region of *traJ*, and terminates at a poly-U stretch also within the untranslated region (Mullineaux and Willetts, 1984). The result is an approximately 78 nucleotide long antisense RNA product, which has the potential to form a duplex with TraJ mRNA. Because the antisense RNA overlaps the ribosome binding site (RBS) for TraJ, hybridization of the two RNAs across this region could block or occlude ribosomes accessing the site, and so block translation. Five alleles of *finP* are known, and these are specific for their complementary *traJ* alleles. The main sequence differences between the *finP* alleles have been identified (Finlay *et al.*,1986), and these differences define the allelic specificities for fertility inhibition observed in the F-like plasmids (Willetts and Maule, 1986).

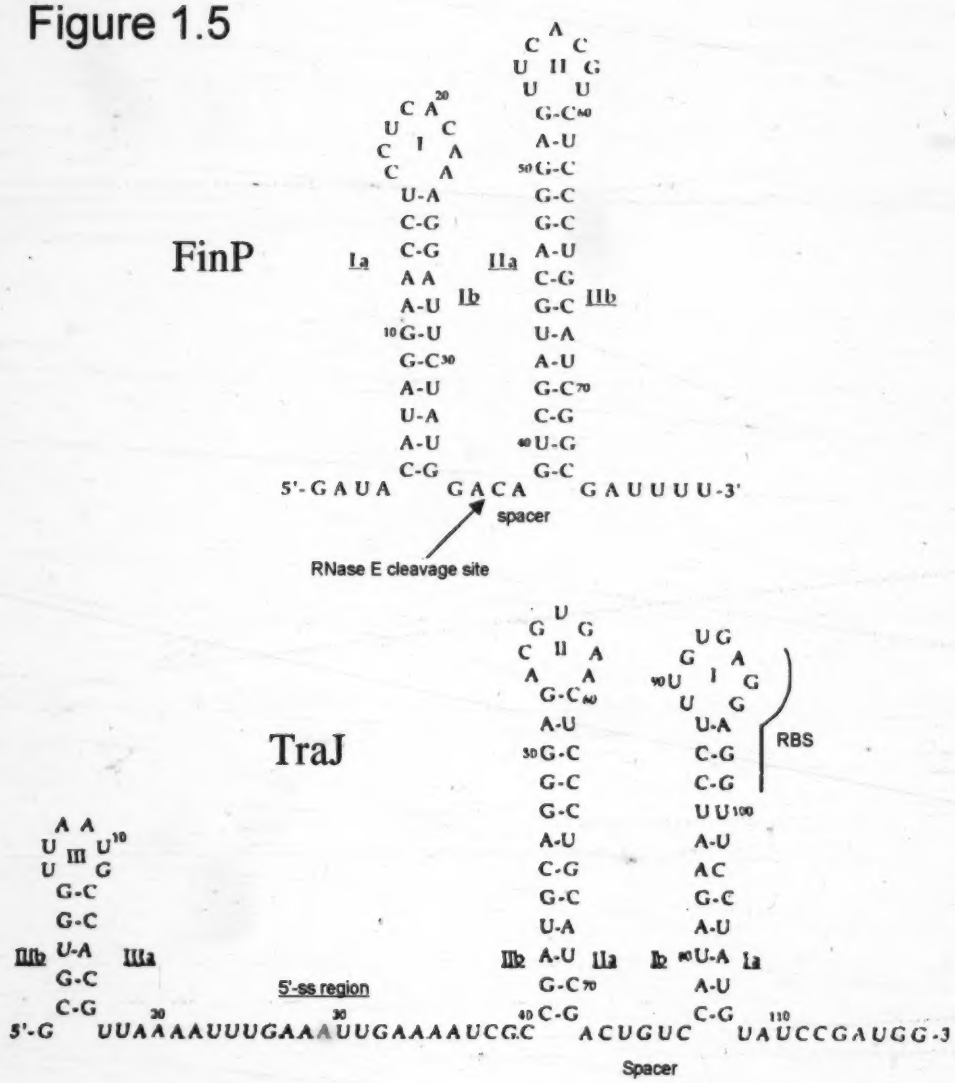
The secondary structure of FinP and TraJ RNAs from F have been defined *in vitro* (van Biesen *et al.*,1993) and are shown in Figure 1.5. Both RNAs have two prominent stem-loop structures; TraJ has an additional minor stem-loop at the 5' end. The sequence differences between the FinP alleles correspond to the loop portions of the stem-loop structures, while the stem structures are highly conserved (Finlay *et al.*,1986). FinP has a relatively short single-stranded region at the 5' end and a very short single-stranded region at the 3' end that ends in a run of Us. This stretch of Us immediately downstream of stem-loop II likely corresponds to a rho-independent terminator. The 4 nucleotide spacer between the stem structures is an important site for RNase E-mediated nuclease activity against FinP RNA (Jerome, unpublished). FinP forms a hybrid with TraJ mRNA *in vitro*, as demonstrated by gel shift assays (van Biesen and Frost, 1994). The overall rate of duplex formation (k_{app}) between FinP and TraJ RNA of 5.0×10^4 $M^{-1}s^{-1}$ is comparable to the rates reported for other sense-antisense RNA systems. The CopA/CopT, RNA I/RNA II, and RNA IN/RNA OUT sense-antisense pairs (Kittle *et al.*,1989; Persson *et al.*,1988; Tomizawa, 1990) are discussed briefly here, and in greater detail in section D. The mechanism of initial hybridization via the stem-loop structures is also consistent with these systems. It was proposed and shown by the lab of Tomizawa that the complementary bases of stem-loop structures in RNA I and RNA II from the ColE1 replication system could form short hybrids in a loop-specific manner (Tomizawa and Itoh, 1981; Masakuta and Tomizawa, 1986). This loop-hybrid complex was termed the "kissing complex". The general reaction can be expressed by the equation:

(1) $$A + T \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} A{:}T$$

where A is the anti-transcript and T is the target transcript. A:T represents the kissing complex and k_1 is the rate constant for the formation of this complex. The constant k_{-1} represents the rate at which the kissing complex dissociates. Once the

Figure 1.5 The 2° structure of FinP and TraJ (117 nucleotides) RNA species. The FinP molecule contains two stem structures, topped by short loops. The stems are separated by a 4 base spacer region which is cleaved specifically by RNase E. The TraJ molecule has a longer 5' single-stranded region and a small 5' stem-loop structure. Stem loops I and II are complementary to those of FinP. The TraJ ribosome binding site (RBS) is located within the top of the stem-loop structure of stem-loop I. Adapted from van Biesen *et al* (1993).

Figure 1.5


FinP
Ia
Ib
IIa
IIb
5'-GAUA GACA GAUUUU-3'
spacer
RNase E cleavage site
TraJ
IIIb
IIIa
IIb
IIa
Ib
Ia
RBS
5'-ss region
5'-G UUAAAAUUUGAAAUUGAAAAUCGC ACUGUC UAUCCGAUGG-3
Spacer

two RNA species have kissed, they can form additional regions of duplex RNA between available single-stranded regions 5' and 3' to the kissing loops. Because extended regions of double stranded RNA are energetically favorable to stem-loop structures, many antisense systems undergo helix propagation along their entire lengths. This additional reaction can involve a number of intermediates. A simplified equation can be expressed as follows:

(2) $$A + T \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} A{:}T \xrightarrow{k_2} AT$$

where AT is the fully duplexed RNA and k_2 is the rate constant for complete nucleation. The rate constant k_{-2} is of such a low magnitude that it is not considered physiologically relevant. In the ColE1 system the rate-limiting step during duplex formation between RNA I and RNA II is the formation of the stable kissing complex (Eguchi *et al.*,1991). Because k_2 greatly exceeds k_{-1} in this system, nearly all the A:T intermediates will rapidly proceed to the AT complex. However, in the CopT/CopA sense-antisense RNA system (plasmid R1) the rate of k_2 was shown to depend upon the length of the single-stranded regions bordering the kissing stem-loop (Persson *et al.*,1990b). A mutant of CopA, designated CopI, had most of the single-stranded 5' leader sequence removed from the base of the second stem-loop. The removal of the leader resulted in more than a 3 log drop in k_2, and so k_2 approached k_{-1} (Persson *et al.*,1990b; Nordstrom and Wagner, 1994). The CopI molecule retained biological activity however. It was concluded that kissing intermediates were sufficient for biological activity within the CopT/CopA system (Wagner *et al.*,1992). It is possible that, as in the CopI/T system, the formation of a kissing intermediate between the FinP and TraJ molecules is sufficient to confer their biological function.

F-like plasmids are observed to not interfere with the translation of TraJ proteins of co-resident plasmids of differing allelic groups. FinP and TraJ RNAs from

different naturally occurring alleles should be incapable of interacting because the sequence differences in the loops should theoretically disallow kissing-interactions. As such, the *in vivo* observations suggest that the *in vitro* work accurately portrays what is occurring in the cells. Additional work with FinP stem-loop mutants also indicates that only the loop-loop interactions are required to occlude the RBS *in vivo* (Koraimann *et al.*,1991; Koraimann *et al.*,1996). In these experiments the *traJ* gene was truncated to remove the *finP* promoter, and the remainder of the *traJ* gene was fused to the reporter gene, *lacZ*. Wild type or mutant FinP was then supplied in trans. Wild type FinP and stem mutants of FinP repressed translation from the TraJ-LacZ RNA due to efficient occlusion of the RBS. The point mutations in FinP loops resulted in several logs of increased LacZ activity, suggesting that the RBS was not blocked because kissing complexes could not form. If the *finO* gene was also added to the system, repression of TraJ-LacZ translation by the FinP loop mutants could be returned to wild type levels. This suggests that FinO is able to mediate duplex formation between imperfectly matched sense-antisense pairs, though the mechanism by which this is accompished is not known. The same study showed that full repression of *traJ* expression is obtained when wild type *finP* is supplied on multicopy plasmids, even in a FinO$^-$ background. This suggests that FinP is the only component of the FinOP system which is absolutely required for repressing transfer, and that this is done in a gene dosage-dependent fashion (Koraimann *et al.*,1996).

III. FinO, an RNA Binding Protein

The *finO* gene is located next to the *traI* and *traX* genes at the distal end of the *tra* region relative to *traJ* and *finP* (Yoshioka *et al.*,1987; Yoshioka *et al.*,1990). At this time it is not known if the gene is under the control of it's own promoter or under the control of an upstream promoter, such as pY. In the case of the F plasmid the *finO* gene is interrupted by an IS*3* element (Cheah and Skurray, 1986). This disruption of the coding region results in the deregulation of transfer. Besides F, several other naturally occurring IncF plasmids (R100-1, Sugino and

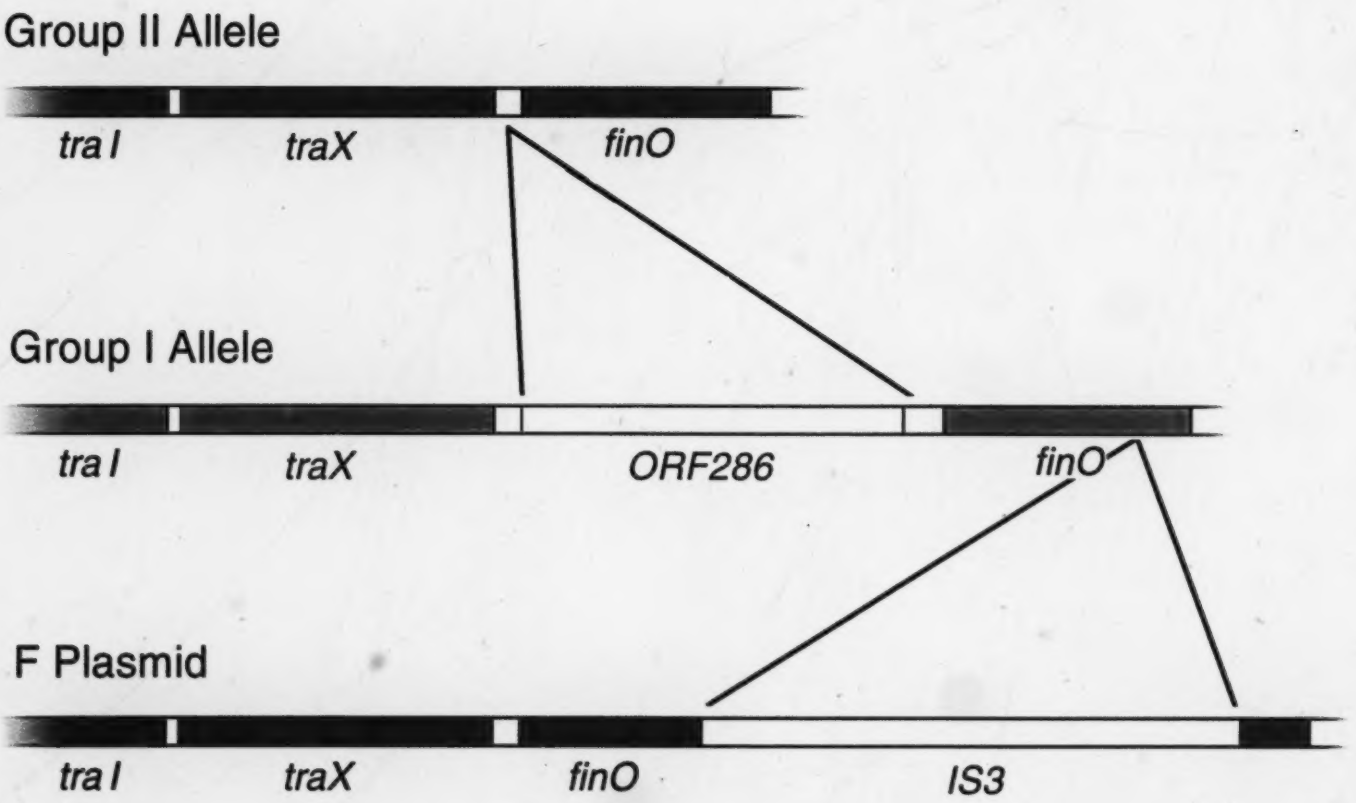
Hirota, 1962; R1-19, Meynell and Datta, 1967) have mutations in the *finO* gene. These plasmids are likewise deregulated for transfer control. Complementation of deregulated IncF plasmids by *finO* from different IncF plasmids indicates that, unlike the *finP* and *traJ* alleles, *finO* is not plasmid-specific. Two alleles of *finO* have been identified (Willetts and Maule, 1986). Group I alleles repress transfer 100–1000 fold while group II alleles repress transfer 20-50 fold. DNA fragments which encode group I *finO* alleles have a neighboring upstream open reading frame variably named *orf*286 (R6.5) or *orf*C (R100) which is absent in the group II alleles (see Figure 1.6). *In vivo* transcription from *orf*286 into the *finO* coding region results in a stable *finO* mRNA. In the absence of *orf*286, *finO* transcripts were too unstable to be detected suggesting that *orf286* acts at the RNA level to increase *finO* mRNA stability in *cis* (van Biesen and Frost, 1992).

finO encodes a 186 amino acid cytoplasmic protein product with an apparent molecular weight of 21 kDa (McIntire and Dempsey, 1987). GST-FinO fusions, which were easier to purify than wild type FinO, are known to bind either FinP or TraJ RNAs *in vitro* with a K_a of 4.0×10^5 M^{-1} and with moderate specificity (van Biesen and Frost, 1994). Binding affinity is strongest for the stem-loop II structure of FinP (van Biesen and Frost, 1994; Jerome, unpublished). Because stem-loop II of FinP/TraJ lacks mismatches or bulges, it was suggested that FinO either refolds the target RNA to allow for the recognition of a specific sequence in the second stem-loop, or that it preferentially binds to uninterrupted, fully duplexed RNA (van Biesen and Frost, 1994). It is now known that high affinity recognition of FinP by FinO requires the presence of a 3' single-stranded tail (Jerome, unpublished), although it is still unknown whether FinO recognizes FinP in a sequence dependent or structurally dependent manner.

Natural FinO protein, when expressed from a *lac* promoter, acts to increase the half-life of FinP RNA from 3 minutes to more than 40 minutes *in vivo* (Lee *et al.*,1992). A fusion protein, GST-FinO was shown to protect FinP RNA from

Figure 1.6 Alleles of *finO*. Group I *finO* alleles contain an upstream open reading frame, *orf286*, which is not present in group II *finO* alleles. Transcription of the *orf286-finO* operon results in a stabilized *finO* transcript, and consequently a higher degree of repression of transfer. F plasmid belongs to group II, but the *finO* gene is interrupted by an inserted *IS*3 element, resulting in the deregulation of transfer.

Figure 1.6


Group II Allele
tra I
traX
finO
Group I Allele
tra I
traX
ORF286
finO
F Plasmid
tra I
traX
finO
IS3

degradation by RNase E *in vivo* and *in vitro* (Jerome, unpublished). FinO also acts to increase the k_{app} between FinP and TraJ RNAs 5-fold *in vitro* (van Biesen and Frost, 1994). As such, FinO appears to decrease translation of TraJ by two mechanisms. First, FinO protein increases the half-life, and consequently the cellular concentration of the antisense RNA. This results in the titration of free TraJ RNA by FinP. Second, FinO increases the rate of duplex formation, thereby increasing the efficiency of titration. The FinO protein is able to bind either free or complexed FinP/TraJ RNAs, as observed by van Biesen and Frost (1994). This differs from the RNA binding protein, Rom, which is only able to bind its target sense-antisense RNAs following the formation of the kissing complex (Tomizawa and Som, 1984).

IV. Interaction of FinO with FisO RNA

One fertility inhibition mutant of *finP*, named *fisO* (Finnegan and Willetts, 1971), produces a particularly unstable transcript *in vivo*. The mutant RNA, FisP contains a G10:U30 mismatch, is bound by FinO at approximately the same rate as wild type FinP RNA and forms duplexes with TraJ at wild type levels *in vitro* (van Biesen, 1994). Despite this, *fisO* does not decrease intracellular levels of *traJ* mRNA significantly when co-expressed with FinO *in vivo* (van Biesen, 1994), suggesting that FisP is degraded *in vivo*, regardless of the presence of FinO protein.

C. Regulation of Transfer in R100 and R100-1

The R100 plasmid, alternatively known as NR1 or R222 is a 94.5 kb plasmid, and belongs to the IncFII incompatability group, (Datta, 1975). It was considered to be of some medical importance, as it was first isolated from *Shigella flexneri* (Nakaya *et al.*, 1960) and carried multiple antibiotic resistance markers. R100 was found to be stably maintained in Hfr (High frequency of recombination) strains. Furthermore, the chromosomally integrated F did not effect the mating efficiency

of R100. However the reverse was not true, R100 could significantly reduce the frequency of transfer of Hfr-associated markers (Nakaya *et al.,* 1960). Likewise F transfer was repressed by co-resident R100 (Watanabe and Fukasawa, 1961). A mutant of R100, named R100-1 was reported to undergo conjugative transfer 10^4 times more frequently than R100 (Sugino and Hirota, 1962), and this derepressed nature was due to a single base insertion in the *fin*O gene of R100-1 (Yoshioka *et al.,* 1987). As a result of this insertion, co-resident R100-1 does not repress transfer events by F^+ or Hfr strains.

It is interesting that the R100 *tra* region shares at least 85% homology with F, as determined by electron microscopic heteroduplex analysis and direct sequence comparisons (Sharp *et al.,*1973; Klimke *et al.,*1997). Even with this degree of similarity between the two plasmids, several interesting differences between their mechanisms of transfer control exist, as noted below.

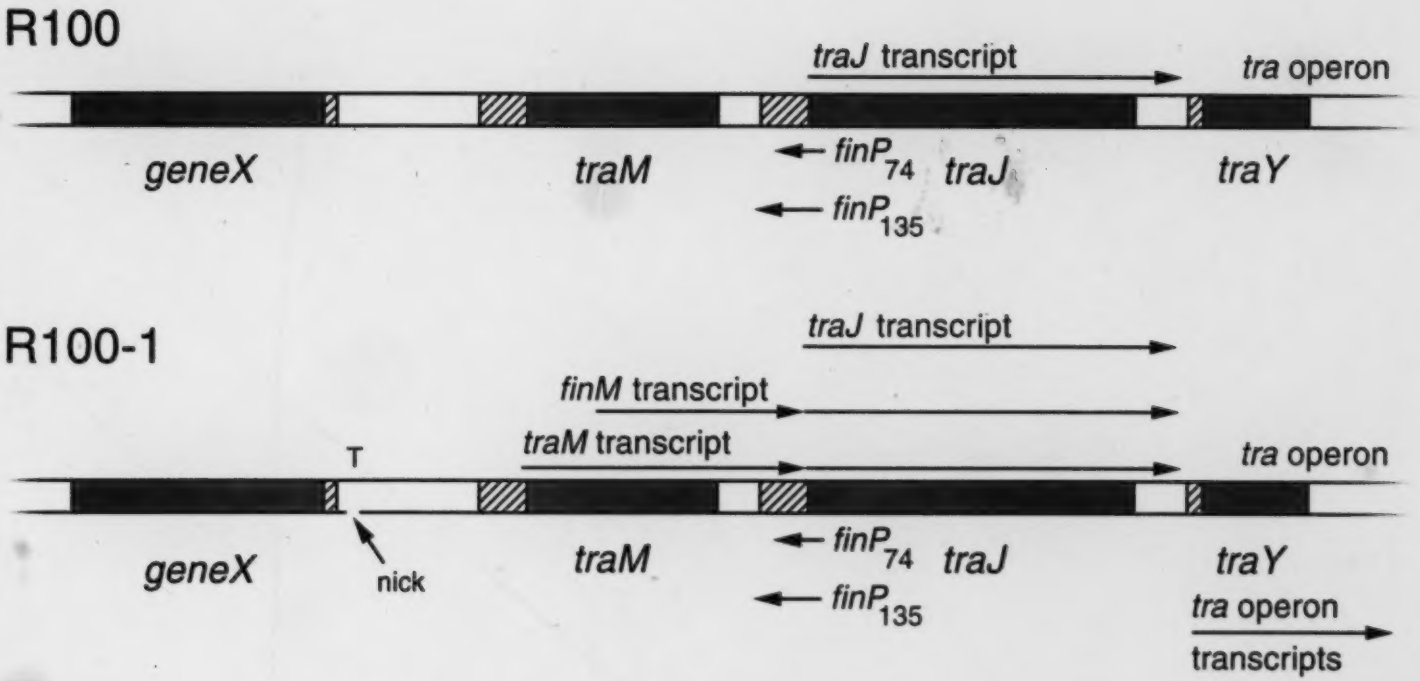
I. R100 FinP and TraJ RNA

Unlike F, R100 and R100-1 both appear to produce two FinP RNAs of differing lengths (Figure 1.7). Though the exact sizes, as determined by Northern analyses and nuclease protection experiments, have varied in the reported literature, the transcripts are considered to be roughly 74 and 135 bases in length (Dempsey 1994; cf. Dempsey 1987). In both R100 and R100-1 backgrounds the 135 base transcript was found at relatively low levels as compared to the 74 base transcript (5% and 95% respectively). It was suggested that in R100/R100-1 two FinP RNAs exist due to transcription past the rho-independent terminator, and/or the action of RNases upon the longer FinP RNA (Dempsey, 1987; Dempsey 1994). Hopefully future work will address this by testing for an accumulation of the long FinP RNA in various RNase mutants.

TraJ mRNA is constitutively transcribed in R100 and R100-1, but the size of the full length transcript has not been determined because the strains tested

Figure 1.7 Organization of the *tra* regulatory region of R100 and R100-1. In R100 two *finP* transcripts are synthesized, one of 74 nucleotides length and another of 135 nucleotides length. The presence of FinO results in the stabilization of *finP* RNA, and so the system is repressed for transfer. In R100-1 FinO is not synthesized, and so expression of TraJ protein is deregulated. TraY protein is synthesized and promotes transcription from the *traM* and *finM* promoters. Some readthrough at the *traM* terminator occurs, such that both long and short *traM* and *finM* transcripts are synthesized. It is uncertain whether long *traM* and *finM* transcripts can form hybrid molecules with *finP* RNA. Adapted from Dempsey (1994).

Figure 1.7


R100
traJ transcript
tra operon
geneX
traM
finP74
traJ
traY
finP135
R100-1
traJ transcript
finM transcript
traM transcript
T
tra operon
geneX
nick
traM
finP74
traJ
traY
finP135
tra operon
transcripts

accumulate transcripts of varying lengths, due either to differing termination (Dempsey, 1987), or elevated RNase activities (Dempsey, 1993). Regardless of the mRNA length, the protein product is 223 amino acids in length, and has a predicted molecular weight of 26 kDa (Inamoto *et al.*,1988).

II. TraM and FinM Transcripts as Anti-antisense RNAs

The *traM* region of the R100 transfer region has been shown to contain three nested open reading frames (Fee and Dempsey, 1986) which putatively encode proteins of 14.5, 10.0, and 8.5 kDa. To date only a 10.5 kDa TraM protein has been identified in cells carrying the R100 plasmid (Dempsey, 1989; Dempsey and Fee, 1990). In plasmid R1, *traM* is expressed from two nested open reading frames (Schwab *et al.*, 1991; Koronakis *et al.*,1985).

Dempsey has shown that the R100 *traM* region contains two promoters (Dempsey, 1989). These are designated the *traM* promoter, which begins transcription upstream of the RBS and the *finM* promoter which is located within the coding region. The F plasmid, in comparison, has two promoters for *traM*, but both are located upstream of the coding region (Thompson and Taylor, 1982; Penfold *et al.*,1996). Unlike F *traM* transcription, which clearly terminates between *traM* and *traJ*, transcription of R100 *traM* and *finM* is known to terminate within the *traJ* leader region. As a result the R100 *traM/finM* transcripts share sequence identity with the 5' end of the TraJ mRNA. It was suggested by Dempsey that the TraM and FinM RNAs might act as anti-antisense RNAs, capable of forming duplex or kissing complexes with the antisense FinP RNA (Dempsey, 1989). This would result in a titration of FinP RNA and presumably some degree of derepression of mating. However, at this time neither the duplex formation between the *traM/finM* transcripts and FinP RNA, nor the subsequent derepression of mating by these transcripts have been irrefutably demonstrated. Finally, transcripts from the *traM* and *finM* promoters frequently bypass the transcriptional terminator located within the TraJ leader region (Figure

1.7), resulting in continued transcription to the end of the *traJ* gene (Dempsey, 1994). Though it has not been demonstrated, it seems possible that TraJ protein could be translated from these extended TraM/FinM transcripts.

III. The Role of FinO in R100

Several models have been proposed by Dempsey to explain the mechanisms underlying the R100 FinO$^+$ phenotype. The two major models of FinO activity for R100/R100-1 are detailed below.

III. a. FinO as an Antisense RNA

The earliest model was based on the effects of a DNA fragment which included an uninterrupted 558 bp open reading frame corresponding to *finO* upon the efficiency of mating. Several mutants of the *finO* ORF were produced and characterized. The data from these, taken together with sequence homology searches suggested that the regulation of transfer was mediated by an antisense FinO RNA (McIntire and Dempsey, 1987; Dempsey, 1993). One such mutant contained the plasmid pWD58, which had R100 *finO* cloned into pBR322 (McIntire and Dempsey, 1987). The R100 *finO* was interrupted by a 4 bp insertion in the *Xma*I site (at position +45 relative to the translational start site) which resulted in a putative peptide product of 75 amino acids. Another clone, pWD59, was interrupted by a 13 bp deletion from the same *Xma*I site, which was estimated to give a peptide of 63 residues in length. In both cases the peptides shared complete identity with the first 49 amino acids of the wild type FinO protein product. The plasmids were tested for their ability to repress the transfer of F' to recipient cells. Both mutant plasmids were shown to repress at levels comparable with clones carrying the entire wild type *finO* fragment. However plasmids which contained deletions at the 3' end of *finO* were derepressed for transfer (same study). From these observations Dempsey concluded that a FinO protein product was not responsible for the observed repression of transfer (McIntire and Dempsey, 1987), at least not through direct interactions with FinP

and TraJ RNAs (Dempsey, 1993). Instead it was proposed that the molecular product responsible for conveying the FinO$^+$ phenotype was an antisense RNA transcribed from the strand opposite the 5' end of the *finO* gene. This antisense RNA was shown to have a 12 base region of complementarity to 12 bases in a 14 base region of the TraJ mRNA (Dempsey, 1987). The proposed mode of action of anti-FinO RNA was to form a duplex across the 12 base region of complementarity with TraJ RNA, causing a conformational change in the TraJ secondary structure, such that the TraJ mRNA could then form either a kissing complex or complete duplex with FinP RNA. A second RNA product was invoked to explain the loss of FinO$^+$ activity in the 3' deletion mutants. This second antisense transcript is proposed to be complementary to the 3' end of the *finO* open reading frame, and is predicted to be essential for the FinO$^+$ phenotype. Because the transcript lacks a designation in the literature, the product is named FinX for the purposes of this thesis.

While the TraJ RNA/anti-FinO RNA interaction model is initially attractive, potential problems should be noted. First, the region of complementarity is fairly strong for R100 TraJ RNA, but comparisons with the sequences of other *traJ* alleles (see sequences from Finlay *et al.*,1986; Frost *et al.*,1994) show reduced complementarity (8 to 9 bases) with anti-FinO RNA. As such, the repression by anti-FinO RNA of all *traJ* alleles is not well addressed. Second, the region of TraJ RNA which is to pair with anti-FinO RNA forms a strong stem loop structure, which would likely occlude more than half of the TraJ site, making it unavailable for the anti-FinO transcript. Third, to date no published results have shown that appreciable levels of the anti-FinO transcript are indeed produced. Finally, the mutation resulting in R100-1 seems to also refute the anti-FinO RNA model. R100-1 has a naturally occurring frameshift mutation due to the insertion of a single A residue at position +55 of the *fin*O gene (Yoshioka *et al.*,1987). The resulting R100-1 FinO translation product terminates at the same location as pWD88, thereby encoding a similarly sized product of 75 amino acids. However, unlike pWD88, R100-1 is completely de-repressed for transfer. To date these

apparent contradictions have not been sufficiently explained to allow a unified theory of the mechanism of the FinO$^+$ phenotype in the R100 plasmid.

III. b. FinO Protein as a Regulator of *tra*M Expression

It was observed that R100 cells contain no detectable TraM or FinM transcripts while R100-1 contained large quantities of both transcripts (Dempsey and Fee, 1990; Dempsey, 1993). Furthermore Dempsey and Fee found that neither the *traJ* gene, nor TraJ protein caused detectable changes in the levels of expression from cloned *traM* and *finM* (Dempsey and Fee, 1990). Because the R100 clones are capable of producing the FinO protein, while R100-1 cannot, it was suggested that FinO protein may be responsible for inhibiting transcription of *traM* and *finM* by binding directly to the region upstream of *traM* as a repressor (Dempsey, 1993). This differs from the F model in which *traM* expression is positively regulated by TraY protein, which is in turn positively regulated by TraJ protein (Penfold, 1995), which is in turn negatively regulated by FinO and FinP (van Biesen and Frost, 1994).

The "FinO as a regulator of *traM* model" was refuted by observations made by Penfold and colleagues (Penfold *et al.* 1996). Plasmid pOX38-Km, which encodes the entire *tra* region of F with the exception of *finO*, was found to produce no *traM* transcripts in cells which carried a co-resident, FinO protein expressing plasmid (pTvB6.11*Nco*I*, van Biesen and Frost, 1992). In cells that expressed reduced levels of FinO (co-resident plasmid pTvB6.11Δ*Sph*I, van Biesen and Frost, 1992), low but detectable levels of the *traM* RNA were synthesized from pOX38-Km. As in the Dempsey experiments, the only real difference between the clones was the presence or absence of the *finO* product, FinO protein. TraM protein levels were then assayed by immunoblots of cells carrying the *finP*$^-$ F plasmid pSLF20 (Lee *et al.* 1992) rather than pOX38-Km. It was found that in the absence of FinP RNA, *traJ* mRNA was translated and *traM* expression ceased to be negatively regulated by FinO expressed from

pTvB6.11*NcoI**. Penfold concluded that expression of FinO protein does act to reduce *traM* expression in certain circumstances, but that this is done via repression of TraJ translation, which in turn prevents TraY expression.

D. Other Sense-Antisense RNA Systems.

Several well characterized sense-antisense RNA systems have been studied to date. Most follow the same basic molecular procedure during RNA duplex formation, but the biological consequences of duplex formation differ from system to system. Replication control of a host of small genetic elements, including the plasmids ColE1, R1, pIP501, and pT181 are due to negative regulation by antisense RNAs. Transposition events by the element Tn*10* are also regulated by an antisense RNA system. In the ColE1 replication control circuit, the antisense RNA causes conformational changes in the sense transcript such that it can no longer act as a primer for plasmid replication (Tomizawa, 1986). In R1, initiation of replication depends upon the accumulation of the RepA protein (Masai and Arai, 1988), and its synthesis is translationally coupled to that of the leader peptide, Tap (Wagner *et al.*,1987). The antisense RNA binds the sense transcript, thereby blocking Tap and subsequently RepA translation (Wu *et al.*,1992). Negative control of copy number in the staphylococcal and streptococcal plasmids pIP501 and pT181 is due to transcriptional attenuation of essential *rep* genes. Antisense binding in these systems causes the sense mRNAs to refold, forming rho-independent terminator structures upstream of the Rep protein coding regions (Novick *et al.*,1989; Brantl *et al.*,1993). Finally, the IS*10*/Tn*10* system involves a unique sense transcript, which does not conform to the usual stem-loop structure found in most of sense/antisense systems (Kittle *et al.*,1989). This sense/antisense interaction results in a direct occlusion of the ribosome binding site for the IS*10* transposase (Ma and Simons, 1990). Many other sense-antisense RNA pairs have been described in prokaryotic and eukaryotic systems. While all antisense RNA systems discovered to date

negatively control biological functions, positive regulation circuits employing antisense transcripts remain a possibility.

I. Replication Control of Plasmid ColE1

The most completely characterized sense-antisense RNA system is that of RNA I/RNAII, which regulates replication of the plasmid ColE1. ColE1 replication requires the presence of a 550 nt RNA primer at the origin of replication to allow for elongation. Synthesis of the transcript, RNA II, into the origin results in the formation of a persistent RNA:DNA hybrid. Subsequent processing of the RNA:DNA hybrid by RNase H produces a 550 nt primer (Tomizawa *et al.*, 1981). The antisense transcript, RNA I, is 108 nt in size, and is complementary to the 5' region of RNA II. Duplex formation between RNA I and RNA II causes a conformational change in RNA II, which blocks the production of the RNase H-sensitive hybrid, thereby blocking DNA synthesis from the origin (Figure 1.8.i; Tomizawa and Itoh, 1981). The timing of the antisense RNA interaction is quite critical. If the transcription of RNA II extends beyond 360 bases then the structure is irreversibly committed to forming the RNA:DNA hybrid at the origin (Tomizawa, 1986). If however RNA I binds RNA II prior to the transcription beyond nt 360, the resulting conformational changes will prohibit the formation of the persistent hybrid. The initial interaction between RNA I and RNA II involves the first three loops of each RNA (see Figure 1.8.ii). Initially a few of the nucleotides in the loops hybridize, resulting in an unstable kissing intermediate (Masakuta and Tomizawa, 1986). Spontaneous dissociation does occur, and this is expressed as the rate constant k_{-1}. A more stable "deep kissing" complex is formed as more intimate loop-loop hybridization occurs. As many as 7 bases per loop can pair to give this intermediate (Eguchi and Tomizawa, 1991). The dissociation rate constant for this complex, k_{-2}, is much lower than the rate constant k_{-1}. Once the deep kissing complex is formed, the now proximal 5' single-stranded leader of RNA I can hybridize with the complementary 3' single stranded region of RNA II, leading to complete nucleation and a resultant stable

Figure 1.8.i Regulation of primer formation at the ColE1 origin of replication. Transcription of RNA II enters the origin (*ori*) of replication and forms a persistent hybrid with the DNA. Processing by RNase H results in the formation of a primer, which allows for DNA synthesis. Hybridization of the antisense RNA I to RNA II results in conformational changes which block the formation of the primer at the origin. Adapted from Wagner and Simons (1994).

Figure 1.8.ii Molecular basis for the hybridization of RNAI and RNA II. RNA I and RNA II form "kissing" intermediates due to base pairing between the complementary loop structures. "Deep-kissing" complexes are formed as the loops deform, allowing the interaction of a greater number of bases within the loops. Hybridization by sequences 5' and 3' to the stem-loop structures results in additional nucleation, and eventually the formation of fully duplexed RNA. Rom protein acts by binding and stabilizing the kissing complex, essentially decreasing the rate of dissociation, k_{-1}. Adapted from Wagner and Simons (1994).

Figure 1.8.i

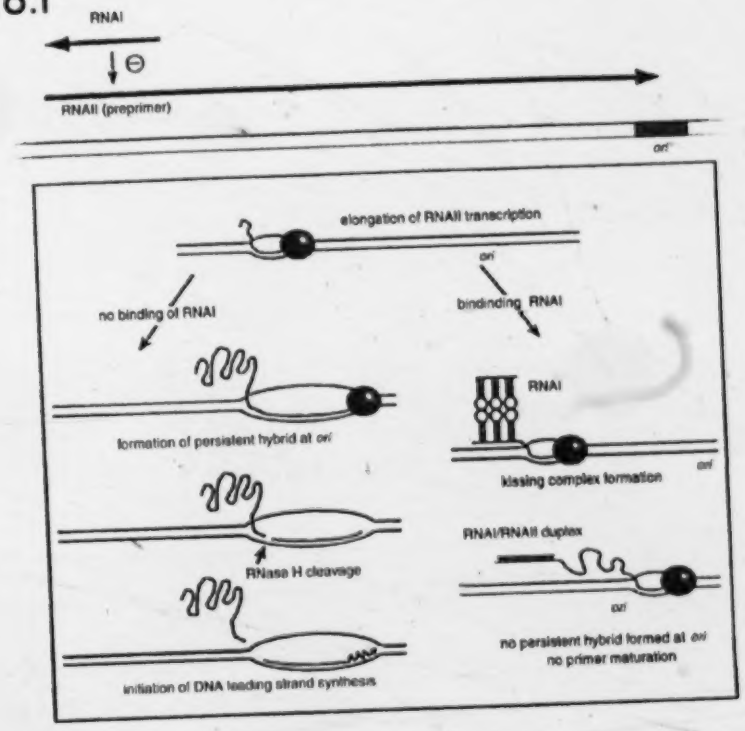
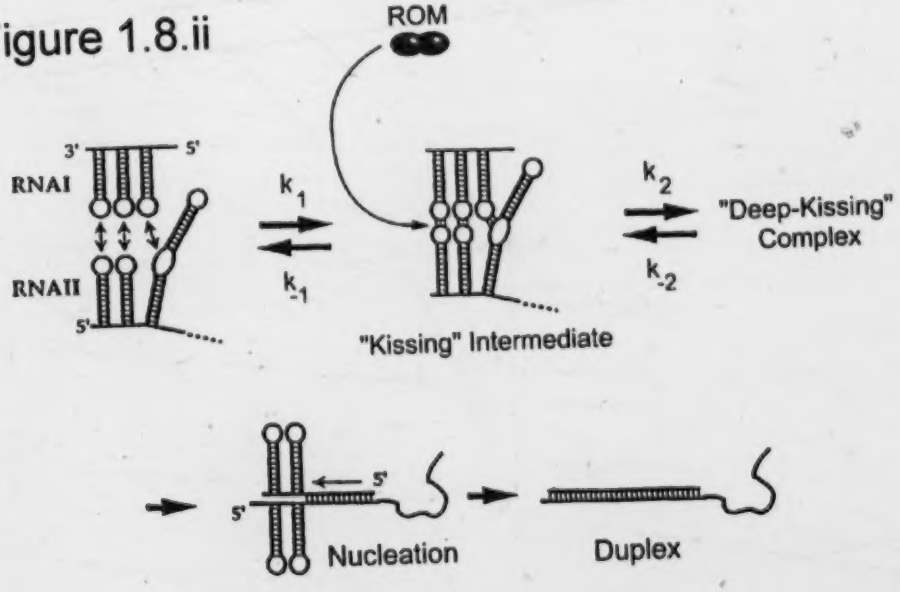


Figure 1.8.ii



duplex. The overall reaction, if we include the intermediate steps, differs from the simplified reaction shown previously. The detailed reaction is as follows:

(3) $$A + T \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} A{:}T^* \underset{k_{-2}}{\overset{k_2}{\longleftrightarrow}} A{:}T^{**} \xrightarrow{k_3} AT$$

where A is the anti-transcript RNA I, T is the transcript RNA II, A:T* is the kissing complex, A:T** is the deep kissing intermediate, and AT is the completely duplexed complex. Because the completely duplexed complex is so stable, the dissociation rate constant k_{-3} is considered negligible. The apparent rate of association (k_{app}) for RNA I and RNA II was calculated to be 7.1 x 10^5 $M^{-1}s^{-1}$ (Tomizawa, 1984).

Because the antisense RNA is transcribed constitutively, and has a relatively short half-life, RNA I concentrations are proportional to the copy number of the plasmid from which they are transcribed (Nordstrom and Wagner, 1994). If the plasmid concentration, i.e. copy number, decreases in the host, the antisense RNA concentration will likewise decrease, and the probability of RNA I: RNA II duplex formation will also decrease. Plasmid replication will occur at a higher rate until the plasmid copy number reaches steady state, at which point primer formation will be efficiently limited by the now more numerous antisense RNA transcripts (Nordstrom and Wagner, 1994). The plasmid encoded protein, Rom (RNA One Modulator), also known as Rop (Repressor of primer), acts by binding and stabilizing the unstable kissing complex (Tomizawa, 1984). This results in a decrease in the equilibrium dissociation constant k_{-1} and so facilitates the formation of the more stable hybridizing complexes (Tomizawa, 1984, 1985, 1990). Functional Rom acts as a dimer with two positively charged α-helices recognizing and binding the RNA, and two negatively charged α-helices acting as "electrostatic rudders" to direct the protein/RNA interactions (Predki *et al.*,1995). Specific mutations in the protein were found to alter the specificity of binding for

the kissing loops of RNA I and RNA II (Predki *et al.*,1995). The effect of Rom upon the RNA I/RNA II duplex formation rate is not as great as the effects seen for other RNA binding proteins and their target RNAs (cf. FinO, van Biesen and Frost 1994; p53, Nedbal *et al.*,1997). Rom effectively doubles the duplex formation rate k_{app} from 7.1 x 10^5 $M^{-1}s^{-1}$ to greater than 10^6 $M^{-1}s^{-1}$ *in vitro* (Tomizawa and Som, 1984), and so roughly halves the copy number *in vivo*.

II. Copy Number Control in IncFII Plasmids.

The IncFII plasmid, R1, utilizes an antisense RNA, named CopA, as its copy control factor. The replicative event at the origin requires an accumulation of RepA protein (Masai *et al.*,1983; Masai and Arai, 1988), which is in turn negatively regulated by CopA. RepA is synthesized from overlapping transcripts, CopT (long) and CopT (short) which are transcribed from different promoters (see Figure 1.9.i). Both CopT transcripts encode RepA as well as the 7 kDa peptide Tap (Translational activator peptide; Wagner *et al.*,1987). Translation of Tap occurs from the CopT transcripts, thereby disrupting an innate 3' stem-loop structure found in the CopT RNAs. Upon translational termination of Tap, the ribosome can then reload to the start codon of RepA (Wu *et al.*,1992). This translational coupling between *tap* and *repA* results in the synthesis of RepA protein and allows for replication of the plasmid. If CopA is present at high concentrations, it will bind CopT transcripts and block translation initiation at *tap*, and so block translational coupling. The *repA* RBS remains occluded by the 3' stem-loop, RepA is not synthesized, and replication does not occur. As in the ColE1 system, the CopA and CopT RNAs form an unstable kissing complex and are believed to also form a stable kissing intermediate. Interactions between the single-stranded spacer regions (located between stem-loop I and stem-loop II) of the RNAs allows for the formation of the stable complex, and subsequent nucleation steps (Figure 1.9.ii; Persson *et al.*,1990a). Deletion of the 5' stem-loop and single-stranded spacer sequences of CopA (mutant CopI; Persson *et al.*,1990b) resulted in a much lower duplex formation rate. Wild-type CopA has a

Figure 1.9.i Regulation of the synthesis of replication factor RepA of plasmid R1. Transcription of CopT RNA occurs from two promoters. Both products encode the RepA product, which is essential for replication of the plasmid. RepA protein synthesis is dependent on the expression of a translationally coupled protein, Tap. Hybridization of CopA RNA to CopT causes repression of *tap* translation. This results in the formation of a stem-loop at the RepA ribosomal binding site (RBS) which blocks ribosomal initiation. Consequently RepA synthesis is blocked, as is plasmid replication. Adapted from Wagner and Simons (1994).

Figure 1.9.ii Molecular basis for the hybridization of CopT and CopA RNAs. CopT and the antisense molecule CopA form a kissing complex due to base pairing between the complementary loop structures. The kissing complex proceeds to a deep-kissing complex as the loops deform to allow a greater degree of hybridization between the nucleotides of the loops. The only region of CopA with significant single-stranded composition is within the spacer region. As a result additional pairing between the molecules depends upon the pairing of the complementary spacer regions. Further nucleation results in the formation of fully duplexed RNA. Adapted from Wagner and Simons (1994).

Figure 1.9.i

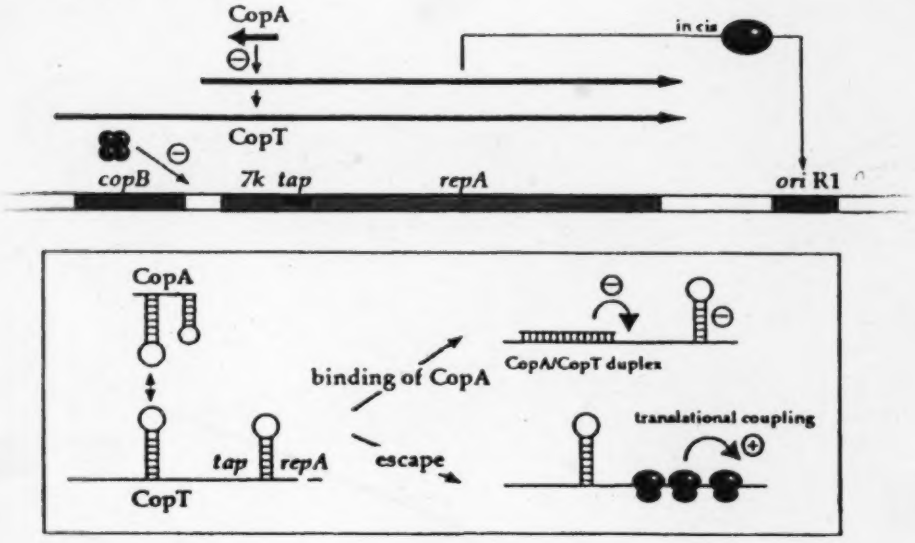
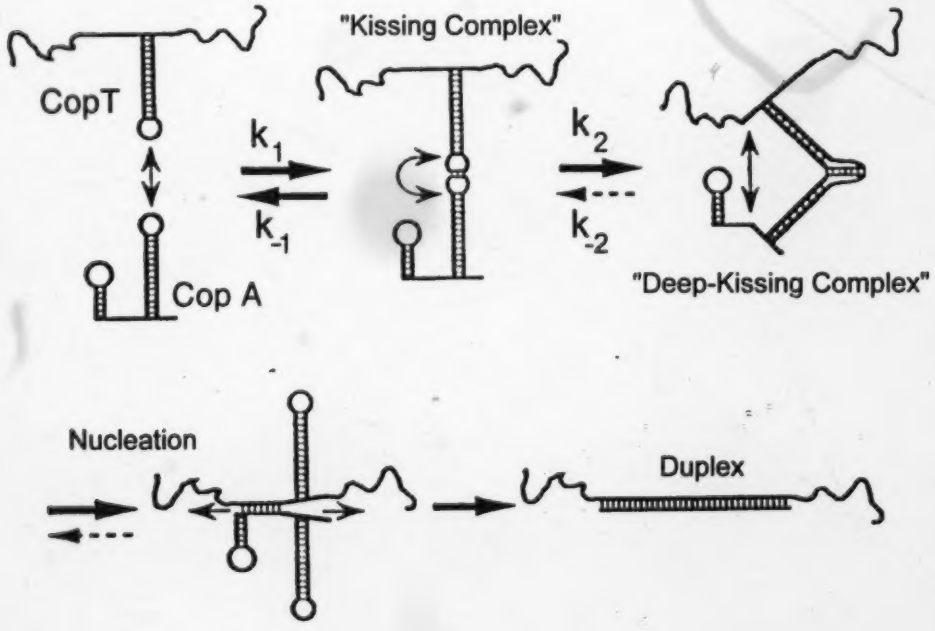


Figure 1.9.ii



duplex formation rate of $>10^{-1}$ s^{-1}, while CopI has a rate of 10^{-4} s^{-1}. Both have the dissociation rate of 10^{-5} s^{-1}. This shows that the deletion of the spacer sequences can have a profound effect on the rate of duplex formation following the initial kissing reaction. The overall k_{app} of 1 x 10^6 $M^{-1}s^{-1}$ for CopA/CopT (Persson *et al.*,1988) is similar to that of RNA I/RNA II. The CopA-CopT complex is cleaved *in vivo* and *in vitro* by RNase III (Blomberg *et al.*,1990). This cleavage more likely acts to clear the accumulating duplex from the cell, rather than reduce RepA synthesis directly, as the kissing complex is sufficient for inhibition of RepA synthesis (Wagner *et al.*,1992).

III. Copy Number Control in pT181 and pIP501

Plasmid copy number control in the staphylococcal plasmid pT181 and the streptococcal plasmid pIP501 depends upon transcriptional attenuation, which is in turn driven by antisense RNAs (Novick *et al.*,1989; Brantl *et al.*,1993). In the pIP501 system, transcription of RNA II can proceed into the coding region of the *repR* gene (Brantl *et al.*,1993). Translation from the nascent mRNA results in the production of RepR protein, and subsequently allows for replication of the plasmid (Brantl and Behnke, 1992). Duplex formation between the sense RNA and the antisense transcript, RNA III, causes refolding of the primary transcript such that a transcriptional terminator stem-loop forms upstream of a poly U domain (Brantl *et al.*,1993). This rho-independent terminator causes premature termination of transcription, and so blocks translation of the RepR protein (Figure 1.10). As in several other systems, complete duplex formation is not absolutely required for the biological activity of the antisense RNA (Wagner *et al.*,1992; Tomizawa, 1990). Binding events which occur prior to full duplex formation are capable of inducing the terminator structure, and so prevent transcription of RNA II into the *repR* coding region (Brantl and Wagner, 1994).

As in pIP501, negative regulation of replication in pT181 depends upon transcriptional attenuation. Unless disrupted by the antisense RNA, transcription

Figure 1.10 Molecular basis for the regulation of the synthesis of replication factor RepR of plasmid pIP501. Short transcripts of RNA II do not encode the RepR protein, while longer transcripts do encode the protein. Transcription of the *repR* gene occurs if the complementary repeats (designated "A" and "a") of RNA II hybridize due to refolding of the molecule. However, if RNA III forms a hybrid with RNA II, the repeats "A" and "B" are sequestered, allowing the formation of the terminator stem-loop. Transcription terminates at the poly-U sequence. Adapted from Brantl and Wagner (1993).

Figure 1.11 Regulation of the synthesis of *IS*10 transposase. The transcript RNA IN encodes the transposase which mediates transposition events of the *IS*10/*Tn*10 element. Hybridization with the stem-loop structured antisense RNA, RNA OUT, results in the occlusion of the *tnp* ribosomal binding site (lower figure). As a result translation is repressed. If RNA OUT does not hybridize to RNA IN translation of transposase occurs, allowing for transposition events. Adapted from Wagner and Simons (1994).

Figure 1.10

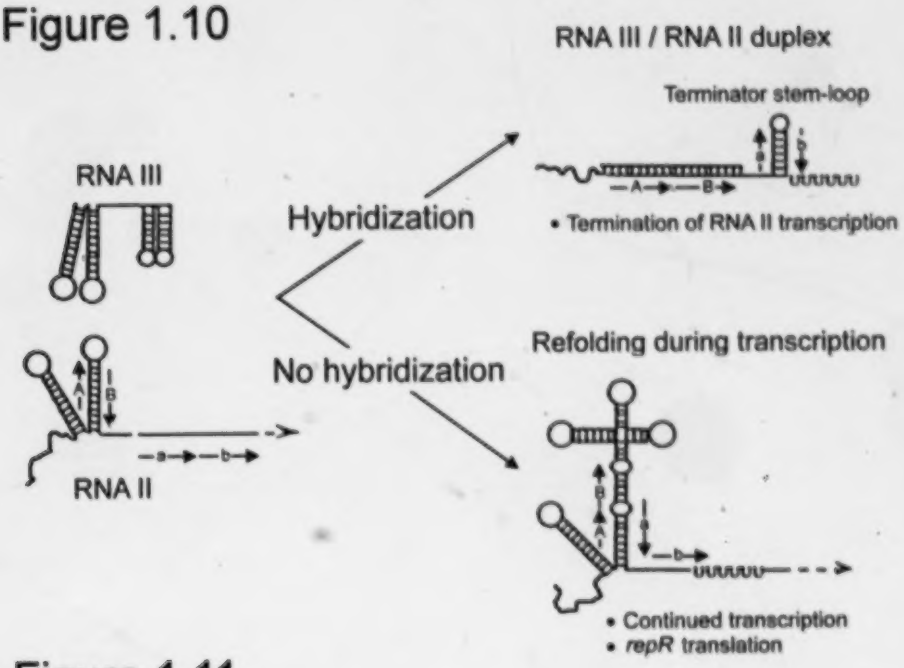
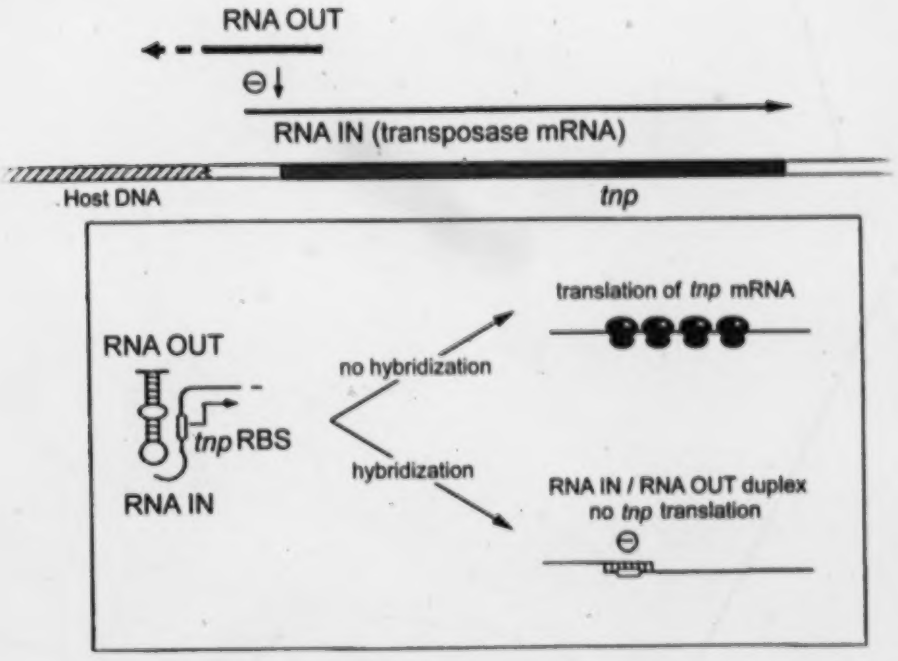


Figure 1.11



of the sense RNA proceeds beyond the untranslated leader region into the *repC* coding region. This results in the translation of the gene product RepC, which is required for plasmid replication. Binding by the antisense RNA to the sense transcript causes the latter to refold, such that a rho-independent terminator is formed (Novick *et al.*,1989). This causes premature termination of transcription upstream of the repC coding region. This essentially limits plasmid replication events.

IV. Control of Transposition Events of the Tn*10*/IS*10* Elements

Transposition events of the tetracycline resistant transposon Tn*10*, and more specifically of the constituent flanking IS*10* sequences is regulated by an antisense RNA, RNA OUT. RNA OUT is transcribed toward the chromosomal DNA from the pOUT promoter (Simons and Kleckner, 1983). The secondary structure of the transcript consists of a single 21 base stem structure, which contains two mismatches. The stem structure is topped by a single-stranded loop domain (Kittle *et al.*,1989; Figure 1.11). RNA OUT has a long half-life of more than 60 minutes *in vivo* (Case *et al.*,1989). Mutations which destabilize the stem structure increase the sensitivity of RNA OUT to exo-ribonucleases, while mutations which lengthen the stem structure increase the degree of sensitivity to the endoribonuclease, RNase III (Case *et al.*,1990), suggesting that the RNA OUT structure is optimized for stability *in vivo*. The sense mRNA, RNA IN is transcribed from within the IS sequence towards, or "into" the transposon, and encodes the IS*10*/Tn*10* gene *tnp* (Kleckner, 1989). The protein product, transposase, is required for the transposition events of the elements (Simons and Kleckner, 1983). RNA IN lacks the secondary structure normally associated with sense/antisense systems. Instead of a complementary stem-loop structure, RNA IN has a 35 base, single-stranded 5' end which is complementary to half of the RNA OUT loop and the 5' side of the stem domain (Kittle *et al.*,1989; see Figure 1.11). Duplex formation is believed to initiate with base pairing between the 5' end of RNA IN and the loop of RNA OUT. Duplex formation then proceeds

along the rest of the complementary sequence, disrupting the stem structure of RNA OUT in the process (Kittle *et al.*,1989). The resulting RNA:RNA duplex is then more susceptible to RNase III degradation, although this is not the main mechanism of control (Case *et al.*,1990). Ma and Simons (1990) demonstrated that the RBS and first codon of *tnp* are located within the first 35 bases of RNA IN, and so could be occluded by duplex formation with the antisense RNA. The authors went on to demonstrate that this occlusion did occur *in vitro*. It is believed that occlusion of the RBS by the antisense RNA directly blocks translation *in vivo* as well, and that this is the major mechanism of negative control blocking the translation of IS*10* transposase (Ma and Simons, 1990). A mutation study of the complementary region of RNA IN and RNA OUT showed that the sequence at nearly every position is critical for maximal inhibition of translation by RNA OUT. The author of the study concluded that the rate of pairing between RNA IN and RNA OUT is nearly optimal *in vivo* (Jain, 1995).

E. RNA Binding Proteins

Many proteins have been identified to date which bind specific RNA species with high affinity and specificity. Several are discussed in detail here. The Rom/Rop protein of plasmid ColE1 is able to bind the RNA I/RNA II pair and increase the stability of the complex, essentially increasing the overall rate of hybridization (Tomizawa and Som, 1984). p53, the eukaryotic tumor suppressor protein has the ability to promote duplex formation between sense and antisense RNA pairs (Oberosler *et al.*,1993). The degree of hybridization promotion between sense-antisense pairs as mediated by p53 is significantly greater than any other RNA binding protein studied to date (Nedbal *et al.*,1997). HIV (Human Immunodeficiency Virus) replication is regulated by two virally encoded proteins, Tat (*trans*-activator protein) and Rev (regulator of virion expression). Both proteins are able to bind target RNA sequences in a specific manner. Tat is a positive regulator for the synthesis of the early mRNAs of HIV, including those

transcripts which encode the regulator proteins (Sodroski *et al.*,1985). As a result Tat positively regulates its own expression. If sufficient Rev accumulates, late mRNAs are preferentially synthesized due to differential splicing (Sodroski *et al.*,1986).

I. Rom/Rop protein of plasmid ColE1

The prokaryotic RNA binding protein Rom (also known as Rop) is known to bind RNA, and to increase the rate of RNA duplex formation between the sense and antisense transcripts RNA I and RNA II (Tomizawa and Som, 1984). Rom binds the target RNAs in a sequence independent, structurally dependent manner (Eguchi and Tomizawa, 1991; Gregorian and Crothers, 1995). Interestingly Rom is incapable of binding the individual sense or antisense RNAs to promote duplex formation. Instead Rom binds and stabilizes pre-existing kissing complexes, thereby promoting duplex formation by decreasing the dissociation rate, k_{-1}. Rom increases the overall k_{app} between RNA I and RNA II 2-fold *in vitro* (Eguchi and Tomizawa, 1990).

The crystal structure of Rom was defined by Banner *et al* (1987), and the protein was further characterized by Predki *et al* (1995) by the alanine scanning method. The protein is a 63 amino acid, which is highly α-helical in nature, and functions as a dimer (Banner *et al.*,1987; Eberle *et al.*,1990). The first α-helical region (helix 1) is a basic domain that binds the target duplex. Replacement of specific residues by alanine in helix 1 reduced or abolished binding to the target RNAs. Mutant Lys3Ala demonstrated reduced binding, while mutants Asn10Ala, Phe14Ala, Gln18Ala and Lys25Ala showed no binding whatsoever (Predki *et al.*,1995). Mutagenesis of Phe14 to tryptophan or tyrosine also negated binding, which is particularily surprising as phenylalanine and tyrosine differ by only one oxygen (Predki *et al.*,1995). Each mutant was examined by circular dichroism (CD) to verify that the alanine substitutions did not cause significant changes to the structure of the α-helical domain. Rather, the loss of affinity was due to the

disruption of binding by specific amino acids to the target RNA (Predki *et al.*,1995). Additionally, this binding is not driven by base-specific interactions, but is dependent upon the duplex structure (Eguchi and Tomizawa, 1991; Gregorian and Crothers, 1995; Chang and Tinoco, 1997). Amino acid substitutions in helix 2, which has a net negative charge, did not affect binding. It is believed that the second α-helix acts as an "electrostatic rudder", which presses the target RNA into the face of the first helical domain (Predki *et al.*,1995).

II. RNA binding and duplex promotion by the protein p53

The tumor suppressing protein, p53 is known to bind DNA in a sequence dependent (Kern *et al.*,1991) or sequence independent manner (Steinmeyer and Deppart, 1988), promote annealing between complementary DNA strands (Oberosler *et al.*,1993) and to promote annealing between sense and antisense RNA pairs (Oberosler *et al.*,1993). The C-terminal amino acids 311 – 393 were shown to promote RNA-RNA hybridization. Alternatively spliced p53 molecules, which do not contain residues 364 – 390, were discovered to be incapable of promoting duplex formation between RNA species (Wu *et al.*,1995). p53 promotes duplex formation between many sense-antisense pairs, and the increase in k_{app} of 480 to 1600-fold is much higher than the increases reported for FinO or Rom protein (Nedbal *et al.*,1997). RNA binding activities and annealing activities were shown to be mediated by two different regions of the protein. p53 is able to bind RNA fragments as short as 20 nucleotides as a monomer. Additional binding to larger transcripts occurs cooperatively (Nedbal *et al.*,1997). The protein has been shown to form dimers and tetramers (Friedman *et al.*, 1993), however it is unknown if a multimer form is required for the RNA annealing functions. At this time the role of specific residues of the protein in RNA binding and duplex formation activities have not been determined. Likewise the RNA structural requirements for binding by p53 are unknown.

III. Tat protein and TAR RNA of HIV-1 and HIV-2

The Tat protein acts by binding the RNA sequence, TAR (trans-activating responsive region) which was originally identified as a 59 nucleotide sequence (Rosen *et al.*,1985). The TAR sequence is located downstream of the transcriptional initiation site, and is present at the 5' end of all HIV transcripts (Rosen *et al.*,1985). Further studies defined the minimal TAR sequence as being at positions +18 to +44 (Jakobovits *et al.*,1988). In the absence of Tat protein, RNA polymerases terminate transcription prematurely (Laspia *et al.*,1989), suggesting that Tat could act to increase transcriptional initiation, stabilize elongation, or both (Sodroski *et al.*,1986). More recent observations suggest that Tat acts specifically to stabilize the elongation process (Karn *et al.*,1994). It has been further suggested that the Tat system may be analogous to the λN protein anti-termination system (Selby and Peterlin, 1990).

The secondary structure of the TAR element has been described as a stem-structure, which is topped by a 6 base loop (Figure 1.12). The stem structure is interrupted by a three base bulge: the region between the bulge and the loop is referred to as the upper stem, and the region below the bulge is the lower stem (Wang and Rana, 1996). The uracil at position 23 (U23) is critical for Tat recognition and binding (Berkhout and Jeang, 1989; Roy *et al.*,1990). Likewise the base pairs G26:C39 and A27:U38 of the upper stem were found to be critical for Tat recognition (Weeks and Crothers, 1991; Churcher *et al.*, 1993). Ethylation of the phosphates between bases G21/A22 and A22/U23 reduced Tat binding significantly, indicating that these phosphates are also important for Tat recognition or binding (Calnan *et al.*,1991b). No other bases or phosphates have been shown to be critical to Tat activity. NMR studies of the unbound TAR sequence did not resolve a definitive structure due to the existence of multiple conformations (Puglisi *et al.*,1992). It was determined however, that the upper and lower stems do exist as helices, and that the bulge is partially stacked between

Figure 1.12 Schematic diagram of the TAR domain from HIV-1 and HIV-2. TAR is composed of a stem structure of 9 canonical base-pairs, which are interrupted by a three base bulge (U23, C24, and U25). The stem structure is topped by a 6 nucleotide loop. Bases critical for Tat protein binding are boxed, and critical phosphates are indicated by small arrows. Sequence from Frankel (1995).

Figure 1.13 Diagram of the domains of RNA binding protein Tat. The molecule is 86 amino acids in length, and contains at least three distinct domains. The N-terminal 49 bases are responsible for stabilizing transcriptional elongation. Residues 49 – 57 constitute the primary binding domain. The arginine-rich domain can be largely substituted by charged residues (+), such as lysine. The natural and minimal amino acid sequences for TAR RNA binding are indicated in standard single-letter code. The C terminal amino acids 59 – 72 constitute a short, glycine-rich α-helical domain which binds the sequence G32, G33, G34 of the TAR loop, and stabilize Tat/TAR interactions. Adapted from Frankel (1995).

Figure 1.12

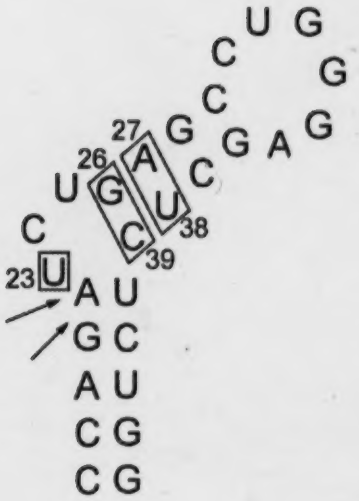
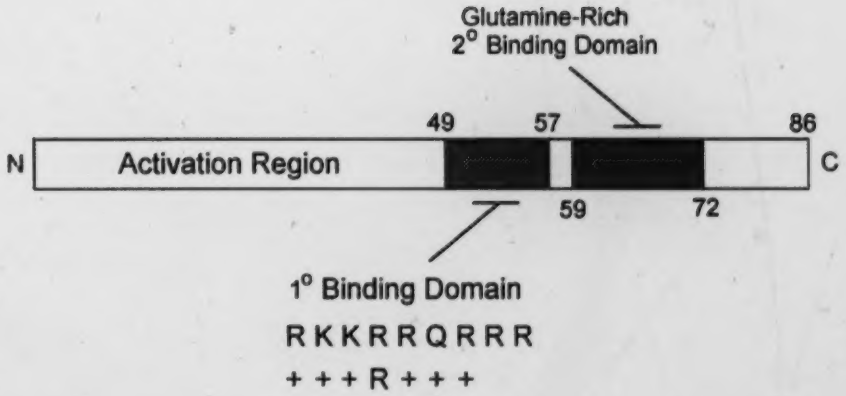


Figure 1.13



the helices (Riordan *et al.*,1992). Overall the TAR bulge introduces a 25° bend to the stem structure under low Mg^{2+} conditions (Zacharias and Hagerman, 1995).

Isolation of the 86 amino acid Tat protein resulted in the co-isolation of a proteolysis product of 38 residues in size, which bound TAR with high affinity and specificity (Weeks *et al.*,1990). The binding domain was further defined as a 9 amino acid, arginine-rich subdomain which corresponds to residues 49 - 57 of the Tat protein (Figure 1.13; Cordingley *et al.*,1990; Calnan *et al.*,1991a). The overall charge density of the subdomain is important to the binding potential of the peptide (Delling *et al.*,1991): substitution of individual arginine residues by lysine does not reduce peptide affinity for TAR (Calnan *et al.*,1991b). The minimum sequence for Tat binding was determined by Calnan *et al* (1991b) to be a single arginine bordered by 3 lysines on both sides. However such a drastic reduction in the size of the peptide causes a drop in binding affinity by 2 orders of magnitude as compared to full-size Tat protein (Tao and Frankel, 1993). The lost affinity was recovered by the reintroduction of the amino acids 57 – 86. The authors of the study suggested that the C-terminal end of the protein is important for stabilizing the binding region (Talanian *et al.*,1990). Free arginine is also able to bind the TAR sequence, but with an affinity 6 orders of magnitude lower than Tat protein (Calnan *et al.*,1991a). Tao and Frankel (1993) suggested that this decrease is due to the loss of electrostatic interactions between the bordering amino acids and the TAR nucleotides. In all cases, whether full Tat protein, or Tat peptides were examined, the binding domain remains unstructured unless bound to the TAR sequence. The short basic domain does become structured following TAR binding, suggesting that the structure of the Tat-TAR complex is defined by the RNA, rather than the protein (Tan and Frankel, 1992; Loret *et al.*,1992; Calnan *et al.*,1991a).

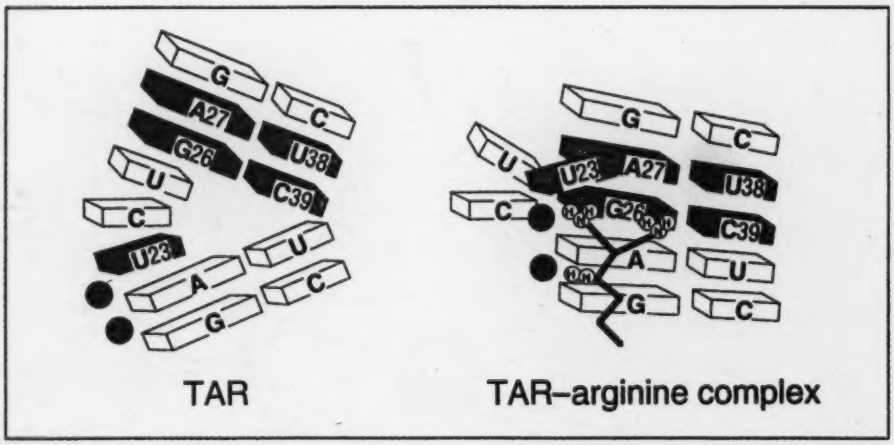
As determined by circular dichroism and NMR spectroscopy, the conformation of the 3 base bulge changes drastically following Tat protein, Tat peptide, or even arginine binding (Tan and Frankel, 1992; Puglisi *et al.*, 1992). These studies

observed that the 3 bases in the bulge are forced out of the stacked conformation between the upper and lower stems, allowing the stems to stack co-axially relative to one-another. Additionally the residue U23 interacts with the base pair A27:U38 to form a base triple. Interruption of the base triple by base substitutions reduces Tat affinity for TAR 8 to 20- fold, depending on the specific mutation (Churcher *et al.*,1993). An arginine from the Arg-rich subdomain interacts with TAR directly via G26 and the critical phosphates (G21/A22 and A22/U23) to give the structure as portrayed in Figure 1.14 (Puglisi *et al.*,1992). Although not previously noted in other RNA-protein complexes, arginine-guanidinium groups have been described to bind guanine residues of DNA in a similar manner (zinc-fingers of DNA binding protein Zif286, Pavletich *et al.*,1991). Karn and associates have also noted that Tat protein recognizes specific bases of TAR via distortions of the major groove near the bulge (Hamy *et al.*,1993). Cross-linking studies confirm that Tat does directly contact U23, U38 and U40 via the major groove (Wang and Rana, 1996). The Tat residue Tyr47, which is immediately upstream of the arginine rich region, binds G26 directly (Liu *et al.*,1996). As a result of Tat binding, the structure of TAR is altered. The ~25° bend noted in the unbound TAR is reduced to <10° due to the relocation of the bases in the trinucleotide bulge (Zacharias and Hagerman, 1995).

In addition to the major arginine-rich subdomain discussed, a second motif for TAR binding has been identified through an *in vivo* study (Loret *et al.*,1992). This 14 residue region (amino acids 59-72) folds as an α-helix, and has a glutamine rich face which binds nucleotides U31, G32, and G33 of the TAR loop directly. This additional binding activity likely accounts for the stabilization of binding by residues 57 – 86 as previously noted (Talanian *et al.*,1990).

Figure 1.14 HIV TAR RNA structure when bound or unbound by arginine residues. The structure is based upon NMR studies of TAR when free or bound by Tat, Tat peptides, or free arginine. Nucleotides and phosphates that are important for Tat protein binding are indicated as darkened boxes and black circles respectively. From Frankel, 1995.

Figure 1.14


G
C
A27
G26
U38
U
C39
C
U23
U
A
C
G
TAR
G
C
U
U23
A27
U38
C
G26
C39
A
U
G
C
TAR–arginine complex

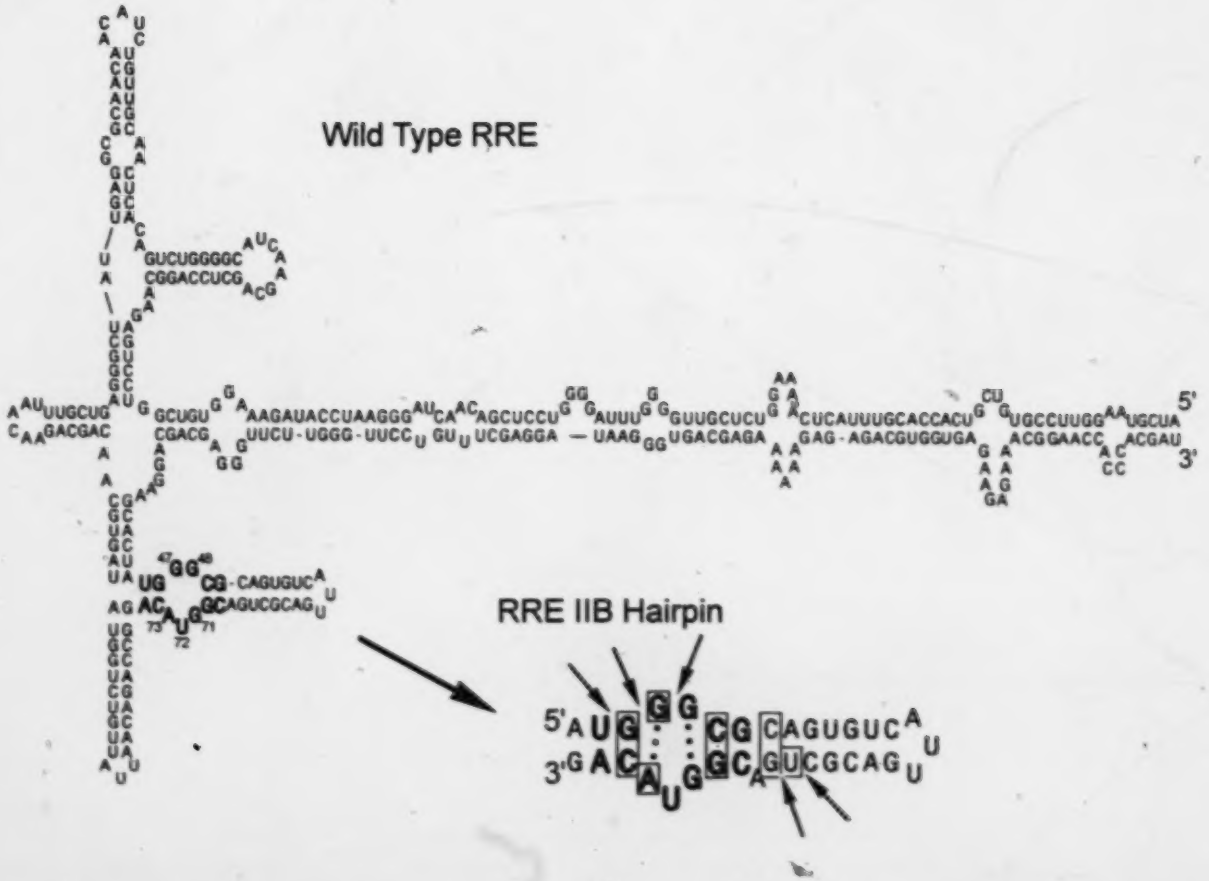
IV. Rev protein and RRE RNA of HIV-1 and HIV-2

Rev protein binds the RNA RRE (Rev Responsive Element) in a specific manner at a number of sites (Zapp and Green, 1989; Daly *et al.*,1989). The binding reaction appears to start with a single Rev protein binding the high affinity site, stem-loop IIB, then additional cooperative binding occurs at the remaining RRE sites (Heaphy *et al.*,1990). The binding by Rev results in the accumulation of differentially spliced, "late" mRNAs such as *env* and *gag*. The binding of Rev to RRE is able to regulate splicing by blocking the entry of the ribonucleoprotein U4/U6.U5 to the spliceosome (Kjems and Sharp, 1993). The ribonucleoprotein U2 is similarly inhibited, while U1 is unaffected by Rev protein (Kjems and Sharp, 1993).

The RRE structure is shown in Figure 1.15. The highest affinity site has been isolated, and is referred to as the purine-rich bubble (Malim *et al.*,1989) or stem-loop IIB (Cook *et al.*,1991). The 'bubble' contains the bulged residue U60 and two non-Watson and Crick base pairs, G35:A61 and G36:G59 (Bartel *et al.*,1991). Both purine:purine base pairs are critical to Rev recognition, although G36:G59 can be replaced by the pair A:A (Bartel *et al.*,1991) or C:A (Werstock *et al.*,1996). Different bases or even an abasic linker can be substituted at U60 without causing a detectable loss of Rev affinity (Heaphy *et al.*,1991), suggesting that position 60 is critical only as a spacer. The base pairs G46:C74 and C49:G70, which are immediately next to the bulge region, are critical for Rev recognition (Heaphy *et al.*,1991). Few substitutions are tolerated at these sites. Though not well characterized, substitutions or chemical modifications of base pair C51:G67 and residue U66 also decrease binding, suggesting that Rev binding/recognition requires more than the internal bulge sequence (Bartel *et al.*,1991; Kjems *et al.*,1992). Several phosphates are also important for Rev recognition (Kjems *et al.*,1992). The five phosphates identified are asymmetrically distributed across stem-loop IIB, suggesting that Rev recognizes the high affinity site in an orientation dependent manner. Like Tat protein, Rev

Figure 1.15 Secondary structure of the RRE RNA of HIV. The high affinity binding site for Rev protein is located at nucleotides 44 to 76, and this region is designated the RRE IIB hairpin. Nucleotides and phosphates that are important for Rev binding are indicated by boxes and arrows respectively. Non-canonical base pairs are indicated by dotted lines (nucleotides G47:A73 and G48:G71) Adapted from Frankel, 1995.

Figure 1.15



has a highly arginine-rich domain which is critical for high specificity binding of the target RNA. However, unlike the unstructured binding domain of Tat, the Rev binding domain is a highly structured, 17 amino acid α-helical moiety (Tan *et al.*,1993). A larger fragment of 40 amino acids, which includes the N-terminal portion of the protein, has a greater affinity and specificity for the RRE, suggesting that the N-terminus is important for conformational integrity of the binding domain (Daly *et al.*,1995). Binding by the protein is dependent upon the non-canonical base pair G48:G71, which causes a local distortion of the sugar-phosphate backbones of both strands. This results in the widening of the major groove of the RNA duplex (Le *et al.*,1994). The Rev α-helical domain may then interact with the RNA via the widened major groove (Williamson *et al.*,1995). Finally, the binding by Rev induces a structural change within stem-loop IIB, changing the duplex from the A form to the B conformation (Auer *et al.*,1994). Base-specific binding by the arginine-rich α-helix was demonstrated by Battiste *et al* (1996). In this study the authors demonstrate that U66 and G67 are directly bound by Arg35, and G70 is directly bound by Arg39. On the other side of the bulge nucleotides U45 and G46 are bound by Arg44, while G47 is bound by Asn40 (Battiste *et al.*,1996). The specific interactions noted are likely via hydrogen bonding rather than electrostatic interactions (Tan and Frankel, 1994) since no substitutions of the nucleotides or the amino acids were tolerated (Battiste *et al.*,1996). Thr34 and several Arg residues (46, 48, 50) appear to be of particular importance for the orientation of the α-helical domain to the major groove. These same residues stabilize the complex via electrostatic interactions with specific phosphates (Tan and Frankel, 1994; Battiste *et al.*,1996).

Following binding by Rev to the IIB structure, additional Rev proteins are recruited to the RRE at lower affinity sites in a cooperative manner (Daly *et al.*,1993a). This multimer formation is required for Rev function *in vivo* (Zapp *et al.*,1991; Huang *et al.*,1991). The wild type Rev protein was shown to form multimers *in vivo* and *in vitro*, while mutants missing amino acids 92 – 112 were unable to form multimers *in vitro* (Daly *et al.*,1993b). Nevertheless these mutant

proteins retained biological activity *in vivo*. This led to the suggestion that Rev polymerization on the RRE may not be necessary for biological activity (Daly *et al.*,1993b). However, further *in vivo* work showed that other regions of the protein were sufficient for multimerization (Madore *et al.*,1994). It is still believed that multimer formation is a prerequisite for Rev activity, and that several regions of the protein are able to mediate Rev-Rev interaction.

F. Summary

Mechanistically the FinP/TraJ antisense system of the F-like plasmids seems to be similar to the RNA I/RNA II and CopA/CopT systems of ColE1 and R1, but with a slightly lower apparent rate of association, k_{app}. FinP/TraJ RNA interactions most likely begin with the formation of a "kissing complex" which would then proceed to a "deep-kissing complex." It is unclear if the FinP/TraJ complex undergoes complete nucleation to give a duplexed product as is seen in both the ColE1 and R1 systems. Functionally however, the FinP/TraJ system is more like the RNA IN/RNA OUT antisense RNA system of IS*10*/Tn*10*. While both systems likely produce double-stranded, RNase III sensitive RNA hybrids, the main mechanism of negative control is the post-transcriptional occlusion of the ribosome binding sites by the antisense RNAs.

The mechanism by which FinO protein interacts with FinP and TraJ RNAs is not entirely analogous to the mechanism by which the Rom protein interacts with RNA I and RNA II. The FinO protein does increase the overall apparent rate of association between the target RNAs, as does the Rom protein. But unlike Rom, FinO also acts by directly binding the individual target RNAs. This allows FinO to protect FinP from nuclease degradation prior to the formation of the kissing complex. As a result, the mechanism of RNA binding by FinO more closely resembles that of the eukaryotic tumor suppressor protein p53 in that they both directly bind the individual sense and antisense targets, and then promote duplex

formation. p53 appears to bind and promote duplex formation across a very broad range of target RNAs. It is not known if FinO can, in a similar manner, promote duplex formation between transcripts other than FinP and TraJ. As is well demonstrated by the RNA binding proteins Tat and Rev, perturbations of the A-form RNA helices are critical for binding by the protein. Without these disruptions, base-specific recognition of the RNA is not possible due to the conformation of the major groove. Recently it has been demonstrated that high affinity binding of FinO to FinP RNA requires the presence of the 3' single-stranded tail (Jerome, unpublished). The tail may play a role in sequence specific protein recognition in a manner similar to the 3 base (single-stranded) bulge of TAR or the unusual 'bubble' found in the RRE stem-loop IIB. Alternatively FinO may recognize FinP and TraJ RNAs in a sequence-independent, structure-dependent manner like Rom.

G. Research Objectives

The purpose of this thesis is to delineate the minimal sequence required for fertility inhibition as promoted by the FinO protein in the F-like plasmids. Previous to these experiments, a 4.0 kb *Pst*I fragment which encodes *finO* was used for expression studies of FinO. The *finO* product has been convincingly shown to confer fertility inhibition, but the possibility remained that downstream sequences may also play a role in fertility inhibition. The experiments presented here evaluate the effects of the downstream sequences upon fertility inhibition by deletion analysis. Deletions into the coding region of the *finO* gene were also made. The resulting *finO* deletion mutants confirm the observations made previously using natural mutants such as R100-1 and F, and add to the body of evidence which contradicts the "FinO as an antisense RNA" model.

The second objective of this research was to determine the functional domains of the FinO protein. *In vitro* studies of the protein have shown that it is able to bind

the target RNAs with moderate specificity, protect the FinP RNA from degradation by RNases, and increase the apparent rate of complex formation between the target RNAs (van Biesen and Frost, 1994). In this work, various fragments of the FinO protein were functionally tested for each of these properties. FinP stability *in vivo* was analysed by mating assays and northern hybridization analysis in the presence of the various fragments of FinO. Protein binding affinities and specificities for the RNA target FinP were examined *in vitro* by gel retardation/migration shift analysis. Finally, the abilities of the FinO derived peptides to promote duplex formation were determined by migration shift analysis under non-denaturing conditions.

Chapter 2. Materials and Methods

Materials and Methods

A. Strains and Plasmids

Genotypes and other data regarding bacterial strains and plasmids used in these studies are listed in Table 2.1 and 2.2 respectively. Plasmids were maintained unless otherwise noted in *E. coli* strain DH5α (*hsdR17 ΔlacU169 φ80 lacZΔM15 recA1 supE44*). β-galactosidase assays were carried out using the Lac$^-$ strain SE5000 (Sm^R *ΔlacU169 flbB3501 deoC1 recA56*). Matings were done using strain ED24 (Spc^R Sm^R *lac*$^-$) as the recipient and MC4100 (Sm^R *ΔlacU169 flbB3501 deoC1*) as the donor organism. pTvB6.11 is a pT7.3 vector containing the 4.0 kb *Pst*I fragment of R6-5 (encoding *traX*, *orf286*, *finO*, *orf86*, and additional downstream sequences) as previously described by van Biesen and Frost (1992). *orf86* was referred to as the *12kDa* open reading frame (van Biesen and Frost, 1992). pOX38-Km, a derivative of the F plasmid, was described by Chandler and Galas (1983). pOX38-Km contains the entire *tra* region from F and the leading region up to and including the RepFIA region. pMCJ184 was constructed by cloning a 327 bp PCR fragment into the *Bam*HI site in the MCS of the plasmid pMC874. The PCR fragment includes the promoter and first 184 bases of the *traJ* coding region, and the entire FinP gene. The primers used were TvB19 (5'TAC CGG ATC CGA AGG TAT CAT CTG AGA TGG AAC) and TvB20 (5'TAC CGG ATC CTG AAT AAC TGC CGT CAG ATT TTC). pMC874 is a translational *lac*-fusion vector, and was originally described by Casadaban and Cohen (1980). pBC-KS+ is a multicopy derivative of pUC-series plasmids which carries the KS multiple cloning site and the chloramphenicol resistance cassette. pBC-KS+ was purchased from Stratagene (1994). pGEX-2T is a translational *gst*-fusion vector, which was purchased from Pharmacia and is described by Frangioni and Neel (1993). pGEX-FO2 was constructed from the parental plasmid pGEX-2T. pGEX-FO2 encodes the GST-FinO fusion protein, and was constructed and described by van Biesen and Frost (1994; and see below).

Table 2.1 List of *Escherichia coli* K-12 strains employed. The genotypes relevant to these studies and references for the strains are listed.

Strain	Relevant genotype	Reference
DH5α	*hsdR17 ΔlacU169 φ80 lacZΔM15 recA1 supE44*	Hanahan, 1983
ED24	*lac*$^-$ Spc^R Sm^R	Achtman *et al.*, 1971
MC4100	*ΔlacU169 flbB3501 deoC1*	Silhavy *et al.*, 1984
SE5000	*ΔlacU169 flbB3501 deoC1 recA56* Sm^R	Silhavy *et al.*, 1984

Table 2.2 List of plasmids employed. The genotypes relevant to these studies, the major promoter and the references for the plasmids are noted. Parental vectors are indicated where applicable.

Plasmid	Relevant Genotype	Promoter	Parental Plasmid	Reference
pACYC184	Tc^R Cm^R	-	-	Chang and Cohen, 1974
pSnO104	Cm^R *finO*$^+$	-	pACYC184	Lee *et al.*, 1993
pBC-KS+	Cm^R	P*lac*	-	Stratagene, 1994
pBC 4.0	Cm^R *finO*$^+$	P*lac*	pBC-KS+	This work
pGEX-2T	Amp^R *gst*$^+$	P*tac*	-	Frangioni and Neel, 1993
pGEX-FO2	Amp^R *gstΩfinO*	P*tac*	pGEX-2T	van Biesen and Frost, 1994
pLJ5-13	Amp^R *finP*$^+$	T7 φ10	pUC19	Jerome, unpublished
pMC874	Km^R *lacZ*$^+$ Δ*Plac*	none	-	Casadaban and Cohen, 1980
pMCJ184	Km^R *traJΩlacZ finP*$^+$	P*traJ*	pMC874	van Biesen, unpublished
pOX38-Km	Km^R F *tra* region	N/A	F	Chandler and Galas, 1983
pT7.3	Amp^R	T7 φ10	-	Tabor and Richardson, 1985
pTvB6.11	Amp^R *finO*$^+$	T7φ10	pT7.3	van Biesen and Frost, 1994

pLJ5-13 was produced by Lori Jerome (Jerome, unpublished). The plasmid carries the *finP* gene downstream of the ϕT7 promoter. The FinP RNA product has the extra sequence GGGGAUC at the 3' end due to the introduction of a *Bam*HI site during cloning. pSnO104 (*finO*$^+$) and the parental (*finO*$^-$) plasmid pACYC184 are described by Lee *et al* (1992) and Chang and Cohen (1974) respectively. pSnO104 carries the entire 4.0 kb *Pst*I fragment of R6-5.

B. Chemicals, Media and Enzymes

IPTG was purchased from Sigma Chemical. Cultures were grown in trypticase soy broth (TSB, Becton Dickinson) + glucose to avoid induction of the *tac* promoter of pGEX plasmids or Luria Bertani broth (LB, Difco) for low level induction. Colonies were grown on Luria Bertani agar or on MacConkey agar (Difco) plus lactose (10g/L). Antibiotic selection was done at the following concentrations; ampicillin (Amp, 50μg/mL), chloramphenicol (Cm, 50 μg/mL), kanamycin (Km, 25 μg/mL), and spectinomycin (Spc, 100 μg/mL). Restriction enzymes and buffers were purchased from Boehringer Mannheim. T7 RNA polymerase and RNase Inhibitor were purchased from Pharmacia. Vent polymerase was purchased from Boehringer Mannheim. Taq polymerase was a kind gift from Dr. Pickard (University of Alberta).

C. Exonuclease III Deletions

The method used for the construction of nested deletions was based on methods previously outlined (United States Biochemical, 1994). All enzymes were obtained from Boehringer Mannheim unless otherwise indicated.

5 μg pBC 4.0 were digested to completion with *Kpn*I and *Hind*III to give 3' and 5' protruding ends respectively. The DNA was subsequently phenol extracted, ethanol precipitated, and washed with 70% ethanol (Sambrook *et al.*, 1989). The

pellet was resuspended and pre-warmed in 50.5 µL 10 X Exonuclease III buffer (660 mM Tris-HCl, pH 8.0, 6.6 mM $MgCl_2$). 250-500 units of Exonuclease III were added, and incubation continued for 10 minutes at 32°C. Larger deletions were obtained when incubated with larger quantities of exonuclease III. 4.5 µL samples were removed every minute, and added to 13.3 µL ice cold S1 nuclease mixture (40 mM potassium acetate pH 4.6, 340 mM NaCl, 1.35 mM $ZnSO_4$, 0.3 units/µL S1 nuclease). Once all samples had been taken, the tubes were incubated at room temperature (RT) for 30 minutes. Digestion was halted by the addition of S1 stop solution (0.3 M Tris-HCl pH 8.0, 50 mM EDTA) and Nuclease S1 was destroyed by incubation at 70°C for 10 minutes. 1.8 µL Klenow solution (20 mM Tris-HCl pH 8.0, 100 mM $MgCl_2$, 0.15 units/µL Klenow) was added with 1.8 µL 125 µM dNTPs to the DNA mixture and incubated for 5 minutes at 37°C to produce blunt ends.

Ligations were done at 16°C overnight following the addition of 40 µL ligation mixture (50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 2.5 µM ATP, 5% PEG, 1 unit T4 DNA ligase). The ligation mixture was ethanol-precipitated and resuspended in 10 µL double distilled water (ddH_2O). 2 µL of ligated product were electroporated into electrocompetent *E. coli* DH5α and grown on TSB plates under selective conditions.

D. Mating Assays

The donor strain *E. coli* MC4100 containing the plasmids pOX38-Km/pBC 4.0, or pOX38-Km/pGEX-FO2 were grown in TSB from a single colony under antibiotic selection to an OD_{600} of 1.0. pOX38-Km/pBC 4.0 and derivatives were grown in TSB + Km, Cm. pOX38-Km/pGEX-FO2 and derivatives were grown in TSB + Km, Amp. The recipient strain ED24 was grown in TSB to an OD_{600} of 1.0 in the absence of antibiotics. 100 µL of each strain were added to 1.0 mL TSB in the absence of antibiotics, and incubated at 37°C for 30 minutes. Mating was interrupted by a 5 second vortex, and cells were diluted 10-fold in ice cold SSC

(150 mM NaCl, 15 mM sodium citrate). 10 µL aliquots were plated from serial dilutions under antibiotic conditions which selected for either donors or transconjugants. The mating efficiency was expressed as the number of transconjugants per donor or as a % mating efficiency as compared to a positive (derepressed) control. Donors were selected by growth on TSB + 25 µg/mL kanamycin, 50 µg/mL ampicillin agar plates and recipients were selected by growth on TSB + 100 µg/mL spectinomycin, 25 µg/mL kanamycin agar plates.

E. β-galactosidase Assays

β-galactosidase activity was determined using *E. coli* SE5000 carrying the plasmids pMCJ184 and derivatives of pBC 4.0. Cultures were grown in TSB to an OD_{600} of 0.8 to 1.2 under antibiotic selection. 50 µL of toluene was added to 3 mL of culture and vortexed 5 minutes. Following a 5 minute settling period at RT, 50 µL of the aqueous phase was added to 1.0 mL of assay solution (6.4 mM KPO_4, 150 mM NaCl, 2mM ONPG (Sigma) , pH7.7) and mixed by inversion. The A_{400} was followed over 10 minutes and the rate of ONPG hydrolysis was determined. Values are expressed in Miller Units:

$$M.U. = 1000 \times [\Delta A_{420} / (t \times vol. \times OD_{600})]$$

Miller, 1972

Where t is equal to the time in minutes, vol. is the volume of cell culture assayed in mL, and OD_{600} is the optical density of the original culture.

F. Preparation of RNA and Northern Transfers

RNA was expressed in *E. coli* strain SE5000 from the plasmid pMCJ184. FinP RNA stability was measured in the presence of GST-fusion proteins as expressed from pGEX-FO2 and derivatives. Where possible, half-lives of the FinP RNA were determined. Cultures were grown in 10 mL TSB from a single colony to an

OD_{600} of 0.8 to 1.0 at 37°C with shaking. Rifampicin was added from a stock concentration of 20 mg/mL in methanol to a final concentration of 200 µg/mL at time zero. Incubations continued at 37°C without shaking. 1 mL samples were removed at the times specified. Samples were immediately microfuged, the supernatants were removed and the cell pellets were flash frozen (-70°C) until sampling was completed. Pellets were then resuspended in 0.2 x volume diethyl pyrocarbonate (DEPC) treated lysis buffer (0.5% SDS, 10 mM Tris pH 7.5, 1 mM EDTA), treated with an equal volume of phenol, and heated at 65°C for 10 minutes. The aqueous layer was extracted and the RNA was ethanol precipitated. The pellet was resuspended in DEPC-ddH_2O and the RNA concentration was estimated at 260 nm.

20-40 µg of RNA were added to 6 X RNA loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 40% w/v sucrose) loaded to an 8% 8M urea polyacrylamide gel (PAG) and electrophoresed at no more than 250 volts. RNA was transferred to Zeta Probe nylon sheets (BioRad) by a Trans-Blot Semi-Dry Transfer Cell (BioRad) and crosslinked at 150 mJoules (GS Gene Linker, BioRad). Prehybridization was done at 37°C for 1.5 hours in 2.5 x SSC, 5 x Denhardt's solution, 1.5% SDS, 50% formamide, 100 µg/mL *E. coli* strain W tRNA, type XX (Sigma). The blot was hybridized 6 hours, then washed in 6 x SSC, 0.1% SDS at 37°C for 15 minutes.

FinP RNA concentrations were determined by the quantification of the band which hybridized the probe TvB17 (5'AAA ATC GCC GAT GCA GGG). In each sample lane the RNA concentration was normalized by quantifying $tRNA_{ser}$, as hybridized by probe SS-2 (5'CCG GTA GAG TTG CCC CTA CTC CGG TTT TAG, He *et al.*, 1993). Quantification of radioactively labeled RNA was done using a phosphorimager (Molecular Dynamics) and the software package, ImageQuant (Molecular Dynamics).

G. Construction of pGEX-Fusion Plasmids

pGEX-FO2 (previously named pFO2, van Biesen and Frost, 1994) and derivatives were constructed by inserting PCR products that contained portions of the *fin*O gene, produced from an R6.5 DNA template (pTvB6.1, van Biesen and Frost, 1992), into the multiple cloning site (MCS) of pGEX-2T (Pharmacia). pGEX-2T is a fusion expression vector which allows for inducible expression from the *tac* promoter of genes that are fused in frame to the 3' end of the Glutathione S-Transferase (GST) gene. pGEX-FO2 was created from a PCR product made using the upstream primer TvB21 (5'ACG GGA TCC ATG ACA GAG CAG AAG CGA CCC G) and the downstream primer TvB11 (5'GGC TTC TCT GGA TCC TTC), both of which contain *BamH*I restriction sites. The fragment was cloned into the *BamH*I site of the pGEX-2T MCS. The protein expressed is the full-sized 186 amino acid (21.2 kDa) FinO protein, fused to the C-terminal end of the 26 kDa GST peptide. The C-terminal deletion mutants pGEX-FO141, pGEX-FO73, and pGEX-FO34 were produced from PCR products made from the upstream primer TvB21 and downstream primers JSA6 (5'GCG AAT TCA GGC ACC GGC TTT CAT GGC), JSA7 (5'GCG AAT TCA CAG TGT GGG CAG GTT CAG), and JSA8 (5'GCG AAT TCA TGG TGG CGT GGT GAC ATT GAT GG) respectively. JSA6, JSA7, and JSA8 all contain a single *EcoR*I restriction site to allow directed cloning into the pGEX-2T MCS. Furthermore, the placement of the *EcoR*I site was engineered to provide an immediate in-frame stop codon. The PCR fragments were cut by *BamH*I and *EcoR*I and cloned into the *BamH*I and *EcoR*I sites of the pGEX-2T MCS. For these plasmids the values 141, 73, and 34 correspond to the number of N-terminal amino acid residues of the FinO protein which are fused in-frame to the C-terminus of the GST peptide.

The N-terminal deletion mutant pGEX-FOC37 was produced from a PCR product made from the upstream primer JSA9 (5'GCG GAT CCG TGA CGG AGC ATA TTT CTC AGG) which contains a *BamH*I site, and the downstream primer JSA4a (5'TTC TGC AGC CTT TAG TGT GAA GGA GG). The fragment was cut by

*BamH*I and cloned into the *BamH*I and *Sma*I sites of the pGEX-2T MCS. For this plasmid, C37 refers to the C-terminal 37 amino acids of the FinO protein which are fused to the C-terminus of the GST peptide.

H. GST-Fusion Protein Purification

This GST-fusion protein purification system is based on the methods of Frangioni and Neel (Frangioni and Neel, 1993) and was described by van Biesen and Frost (1994). Specifically, 100 mL cultures of DH5α/pGEX-FO2 (or derivative) were grown from a single colony in TSB + 100 µg/mL ampicillin at 37°C with shaking. At an OD_{600} of 1.0 to 1.2, the cultures were induced for 5 hours with 0.5 mM IPTG (final concentration), pelleted by centrifugation and frozen at –20°C for storage.

Pellets were resuspended in 5 mg/mL lysozyme dissolved in STE (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA) and incubated for 30 minutes on ice. Dithiothreitol (DTT) was added to a final concentration of 4 mM and N-laurylsarcosine (sarkosyl) was added with vortexing to a final concentration of 1.5%. Lysis was done using a French press (American Instrument Company), at a constant cell pressure of 1300 kPa. Triton X-100 was added to the lysate to a final concentration of 2%, vortexed and incubated for 30 minutes on ice. It was then cleared of insoluble debris by centrifugation for 10 minutes at 10 000 x g. 100 µL of a 50% glutathione-agarose bead slurry (S-linked, Sigma) was added, and binding was allowed to proceed for 2 hours at 4°C. The slurry was washed 3 times with 20 mL TEB (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 100 µg/mL BSA). GST-fusion proteins were eluted by incubation in one volume 20 mM reduced glutathione (Sigma) dissolved in TEB for 10 minutes at RT.

Proteins were stored in 20% glycerol, 1 mM DTT, 200 µg/mL RNase-free BSA (Boehringer Mannheim) at –20°C for up to one month with no detectable loss of activity.

I. *In vitro* Transcriptions

In vitro transcription were done as previously described in van Biesen and Frost (1994) with the following modifications. $TraJ_{184}$ RNA was synthesized from PCR-generated DNA templates using the 5' primer TvB15 (5' TCG AAT TCT AAT ACG ACT CAC TAT AGA CGT GGT TAA TGC CAC G), which encodes a functional T7 promoter and the 3' primer TvB14 (5'CCT GAA TAA CTG CCG TCA G). The resultant RNA product was found to be 184 bases in length rather than 211 as previously reported (van Biesen and Frost, 1994). FinP RNA was synthesized from *BamH*I digested pLJ5-13ϕ10, such that the RNA transcript has a total length of 86 bases. The FinP RNA has the additional sequence of GGGGAUC at the 3' end of FinP due to the *BamH*I site (Jerome, unpublished).

[^{32}P]-labeled RNA was synthesized using α[^{32}P]-UTP (NEN/DuPont, 3000Ci/mmol) while "unlabeled" or [^{3}H]-RNA was synthesized using 5,6-[^{3}H]-UTP (NEN/DuPont). RNA synthesis was carried out using T7 RNA Polymerase for three hours at 37°C in the presence of 0.5 units/μL RNasin. Aliquots were removed prior to DNase treatment for the determination of specific activities.

RNA loading buffer was added to the transcripts, which were then electrophoresed at 250 volts on an 8% 8M urea PAG for 1 to 2 hours. Appropriate bands were visualized by exposure to X-ray film (Kodak X-omat AR), cut out and eluted overnight in DEPC-treated elution buffer (0.5M NH_4OAc, 1mM EDTA) at 37°C. Eluted RNA was treated with an equal volume of phenol. The aqueous phase was removed, ethanol precipitated, and washed with 70% ethanol prior to drying. Dried pellets were resuspended in DEPC-H_2O and frozen at –20°C. [^{3}H]-labeled RNAs were identified for excision from the gels using [^{32}P]-labeled transcripts as markers.

J. Mobility Shift Assays: RNA Binding, Competition Assays

The specified concentrations of purified GST-fusion proteins were individually incubated with 5 x 10^{-10} M [^{32}P] FinP or [^{32}P] $TraJ_{184}$ RNA (15 fmoles) in a total volume of 30 µL binding buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 10 mM NaCl, 100 µg/mL RNase free BSA, 10% glycerol, 0.3 to 0.5 units/µL RNasin) for 30 minutes at RT. Samples were loaded onto an 8% non-denaturing PAG and electrophoresed at 150 volts for 1 hour. No RNA loading buffer was added. Competition assays required the mixing of the two RNA species at 4°C just prior to their addition to the reaction mixture.

K. Duplex Formation Assays

Fusion protein was added to 75 µL duplex buffer (30mM Tris-HCl pH 8.0, 5 mM $Mg(OAc)_2$, 0.5 mM EDTA, 50 mM NaCl, 50 µg/mL RNase free BSA, 0.3 units/µL RNasin) to a final concentration of 1.7 x 10^{-7} M (5.1 pmoles), which corresponds to the protein concentration which gave ~100% binding in the mobility shift assays. [^{32}P]-FinP RNA was mixed with [^{3}H]-$TraJ_{184}$ at 4°C. The RNA mixture was then added to the protein/buffer mixture at time zero to a final concentration of 5.0 x 10^{-10} M [^{32}P]-FinP and 5.0 x 10^{-9} M [^{3}H] $TraJ_{184}$. Aliquots were added to ice cold formamide stopping solution (95% formamide, 20 mM EDTA, 0.05% Bromophenol Blue, 0.05% Xylene cyanol FF) at the times specified. All samples were loaded to an 8% non-denaturing PAG, and electrophoresed at 150 volts for 1 hour.

L. Calculations

I. Calculation of the Association Equilibrium Constant, K_a

The association equilibrium constant (K_a) was calculated as previously described (van Biesen and Frost, 1994, Tsai *et al.*, 1990). Briefly, K_a is calculated from the concentration of GST-fusion protein required to cause a 50% shift of labeled

RNA on the gel. Calculations of the % RNA shifted by the protein fusions GST-FinO141 and GST-FinO73 needed to include values from the quantification of RNA trapped in the wells since these proteins displayed considerable aggregate formation. Additionally the calculated K_a values may be artificially low since we cannot determine what fraction of the protein preparations are active (van Biesen and Frost, 1994). The equilibrium can be defined as:

(4) $$B + O \underset{K_D}{\overset{K_a}{\longleftrightarrow}} BO$$

And so the equation for determining K_a is as follows:

(5) $$[BO] / ([B] \cdot [O_o])$$

Where [B] is the concentration of free FinO binding sites and [BO] is the concentration of bound FinO binding sites. $[O_o]$ is the initial protein concentration in the reaction. When 50% of the target RNA shifts on the gel, [B] equals [BO], giving the formula:

(6) $$K_a = 1 / [O_o]$$

II. Calculation of the Binding Rate Constant, k_{app}

The second-order binding rate constant for FinP/TraJ$_{184}$ duplex formation (k_2, Persson *et al.*, 1988; k_{app}, van Biesen and Frost, 1994) was determined from a log plot of the % free [^{32}P]-FinP RNA remaining over time. The equilibrium is expressed in equation 1.

(1) $$A + T \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} A{:}T$$

Where A is the antisense RNA, FinP, T is the target RNA, $TraJ_{184}$, and A:T is the sense/antisense RNA complex. Because there was a 10-fold excess of [^{3}H]-$TraJ_{184}$ RNA, pseudo-first-order kinetics were obtained. The pseudo-first-order rate constant, k_1' can be calculated from $t_{1/2}$, the time required for 50% of the labeled molecules to become complexed (Persson *et al.*, 1988):

(7) $$k_1' = \ln 2 / t_{1/2}$$

The second-order rate constant can be calculated as:

(8) $$k_{app} = k_1' / [S]$$

Where S is the RNA species in excess.

Chapter 3. Deletion Mutations of the Fertility Inhibition Protein FinO Cause Derepression of Transfer of the F Plasmid.

Introduction

In the F-like plasmids, transfer is regulated by the two component fertility inhibition system, FinOP. The first component, FinP, is an antisense RNA which can hybridize to the target mRNA, TraJ. The formation of the FinP:TraJ RNA complex is believed to occlude the TraJ ribosomal binding site. Without TraJ protein synthesis, the rate of plasmid transfer decreases approximately 1000-fold. The second component of the fertility inhibition system is the *finO* product. Complementation of *finO*$^-$ plasmids results in the recovery of transfer repression. The open reading frame for *finO* is 558 base pairs in size, and has been shown to encode a protein product of 21.2 kDa (van Biesen and Frost, 1992).

The simplest model for the FinO$^+$ phenotype proposes that the FinO protein directly binds FinP RNA, thereby protecting the transcript from RNase activities (van Biesen and Frost, 1994). *In vitro* experiments have shown that the FinO protein is able to bind both FinP and TraJ transcripts (van Biesen and Frost, 1994). Additionally, the protein acts to protect FinP from RNase E activity *in vitro* (Jerome, unpublished). Despite these results two alternative (and complex) models have been invoked to explain the FinO$^+$ phenotype (McIntire and Dempsey, 1987; Dempsey, 1993; see chapter 1).

The purposes of these experiments are three-fold. First, *orf286,* upstream of *finO*, has been shown to play an important role in fertility inhibition (van Biesen and Frost, 1992). In this report the role of sequences downstream of *finO* in fertility inhibition are determined. Second, I wanted to determine if any deletions from the 3' end of *finO* would cause derepression of transfer. Third, the experiments are intended to determine whether the alternative models, as proposed by Dempsey, are valid.

Results

I. Exonuclease Digestion of *finO* Sequences

In these experiments the 4.0 kb *Pst*I fragment containing an R6-5 *finO* was excised from the plasmid pTvB6.11 (van Biesen and Frost, 1992), and cloned into the MCS of the vector pBC-KS+ (Stratagene, 1995). The product of this cloning is named pBC 4.0. The fragment was cloned into the multiple cloning site (MCS) in an orientation opposite that of the *lac* promoter and the chloramphenicol cassette (Figure 3.1). Exonuclease III was then employed to processively remove sequences downstream of *finO*, including the reading frame *orf86* (previously named *orf 12kDa*; van Biesen and Frost, 1992). Digestion of the *Pst*I fragment was encouraged by the production of a 5' overhang (*Hin*dIII cleavage) in the vector, proximal to the fragment. In many instances exonuclease digestion proceeded into the *finO* reading frame, resulting in the synthesis of truncated *finO* products. Vector sequence was protected from exonuclease digestion by prior enzymatic cleavage at the *Kpn*I site in the MCS. The resultant 4 base 3' overhang blocks exonuclease processivity.

Locations of the deletion mutants employed in this study are shown diagrammatically in Figure 3.1. Map locations were determined by either restriction map analysis (Figure 3.2) or by sequencing (see materials and methods). Clones were named on the basis of the position of the truncation, relative to the first base of the *finO* coding region. As such, clones with names corresponding to values less than +557 have deletions extending into the coding region of *finO*. Negative values correspond to deletions extending into either *orf286* or *traX*. The mutant pBC-1650* and others are denoted with an asterisk to indicate that these clones were mapped by restriction analysis only, and so the values given (relative to the *finO* start site) are only approximate.

Figure 3.1 Diagram of map positions of deletion mutants of plasmid pBC 4.0 as produced by exonuclease III. The top line indicates the map position of the genes in 0.5 kb demarcations. Genes from the 4.0 kb *Pst*I fragment of R6-5 are denoted in the second line. Hatched diagonal lines indicate vector sequence. The direction of exonuclease III digestion is denoted by the cartoon enzyme. Derivative plasmids are aligned below the second line relative to the first codon of *finO*, designated +1.

Figure 3.1

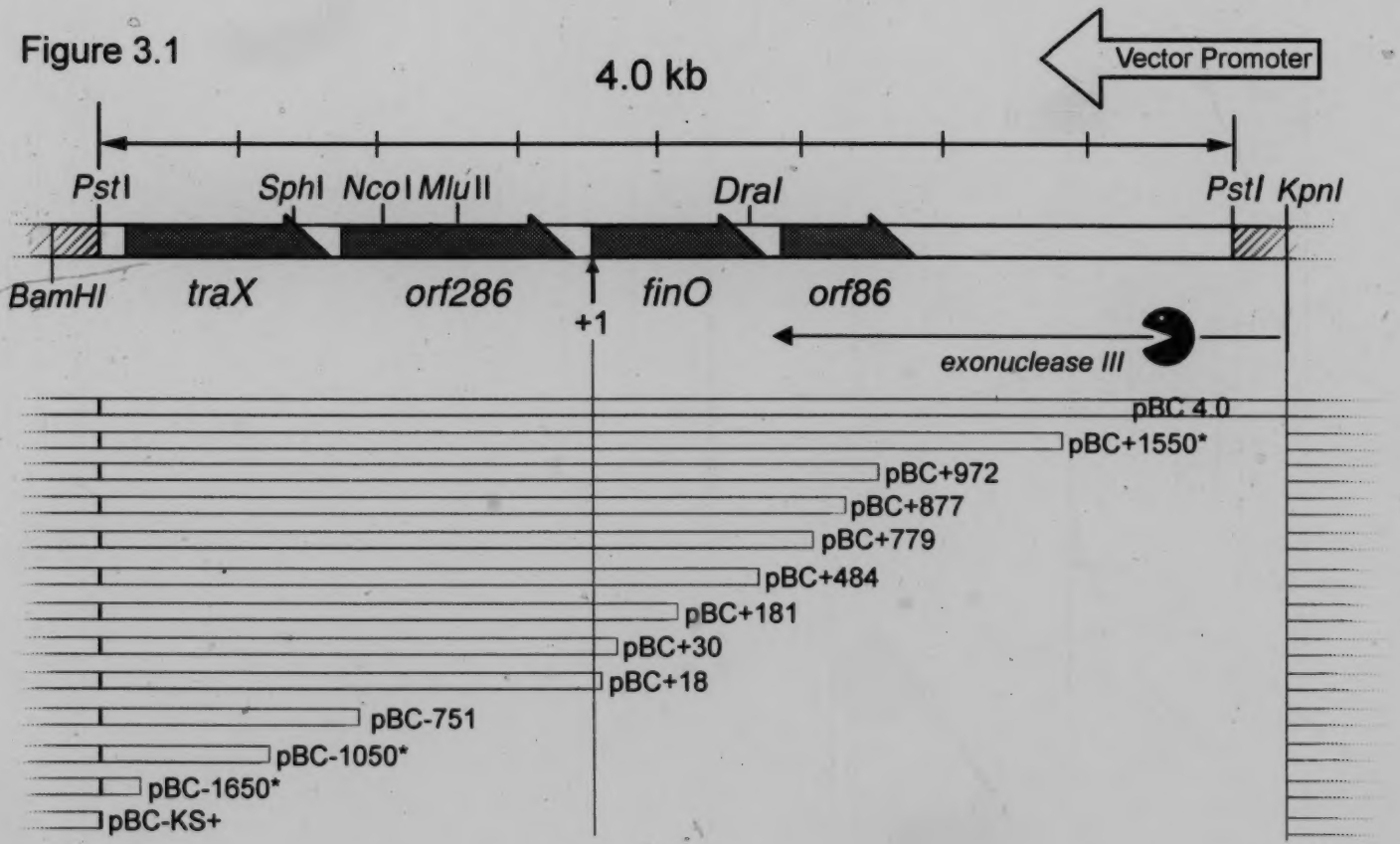

4.0 kb
Vector Promoter
PstI
SphI
NcoI
MluII
DraI
PstI
KpnI
BamHI
traX
orf286
+1
finO
orf86
exonuclease III
pBC 4.0
pBC+1550*
pBC+972
pBC+877
pBC+779
pBC+484
pBC+181
pBC+30
pBC+18
pBC-751
pBC-1050*
pBC-1650*
pBC-KS+

Figure 3.2 Restriction map of the exonuclease III-generated mutants of the plasmid pBC 4.0. pBC 4.0 was digested by *Bam*HI and *Dra*I and resolved on an 8% agarose gel. The four digestion products, named fragments A, B, C, and D, are 3.99, 2.30, 1.37, and 0.34 kb in size respectively (lane3). The truncated plasmids were digested and resolved in an identical manner. Exonuclease digestions proceeded in a directed manner from the vector *Hin*dIII site toward the 4.0 kb inserted sequence, resulting in a decrease of mass of fragment A (lanes 4-14). Mutants in which the exonuclease digestion had crossed the *Dra*I site of *finO* produced three fragments, C, D, and a new hybrid fragment which contained the remainder of the undigested sequence from fragments A and B. Lane 1 contains the DNA molecular marker III (Boeringher Mannheim) with fragment sizes 21.2, 5.1, 4.9, 4.2, 3.5, 2.0, 1.9, 1.5, 1.3, 0.94, 0.83, and 0.56 kb from top to bottom. Lane 2 contains the control vector pBC KS+, digested by *Bam*HI and *Dra*I. Lane 12 contains a cointegrate product. The cointegrate was not used in further studies.

Figure 3.2

II. Qualitative Assays of FinO Activity *in vivo*.

The pBC 4.0 deletion mutants were tested for their ability to protect FinP RNA and promote duplex formation *in vivo*. The construct pMCJ184 encodes a truncated TraJ transcript of 184 nucleotides, which is translationally fused to a complete *lacZ* gene. Full length FinP RNA is also produced from pMCJ184. Both pBC 4.0 (or derivatives) and pMCJ184 were transformed into the *recA- E. coli* strain SE5000, which is *lacU169*. Transformants were plated under selective conditions on MacConkey-lactose agar. Cells which do not carry the *finO* gene, or that encode a non-functional *finO* gene product should be incapable of stabilizing FinP RNA. As a result, the FinP RNA would not effectively regulate translation from the *traJ::lacZ* fusion transcripts. Colonies which arise from these cells would be phenotypically Lac^+, and on MacConkey-lactose plates they would be bright red in color. In the presence of a functional *finO* gene product the FinP RNA should be stable, and so significant repression of translation from the fusion transcript would occur. The resulting colonies would have a Lac- phenotype, and would be white in color. A mutant producing partially functional FinO would be predicted to give a pink or light pink-colored colony. The colony color of each pBC 4.0 derivative when plated on MacConkey-lactose plates (as described above) is given in Table 3.1. The plasmid derivatives of most interest are pBC-751, pBC+484, pBC+779, and pBC+1550*. Colonies containing pBC+1550*, which encodes an intact *orf86*, were white in color. Colonies containing pBC+779, which has a deletion extending well into *orf86*, were also white, suggesting that *orf86* does not play a regulatory role in FinO expression. pBC+484 has the smallest deletion of the *finO* gene of any mutants mapped in these experiments. A 13% deletion at the 3' end of the *finO* gene was sufficient to give red colonies, indicating significant deregulation of the fertility inhibition system. The plasmid pBC+484 is expected to produce a FinO peptide of 161 amino acids. By comparison pBC-751 and pBC-KS+, neither of which encode any portion of *finO*, had equally red colonies upon visual inspection.

Table 3.1 Genotype and phenotypes of deletion derivatives of plasmid pBC 4.0. The number value in the plasmid name corresponds to the number of bases remaining downstream of the first codon of *finO*. The predicted size of the FinO protein product is indicated for each plasmid. The data in the third and fourth columns, color on MacConkey Lactose plates and β-galactosidase activity in Miller Units, were determined from clones that contained the co-resident *traJ-lacZ* fusion plasmid, pMCJ184. The last two columns indicate the number of transconjugant progeny per donor organism and the % mating efficiency relative to the *finO*- plasmid, pBC-KS+. The co-resident conjugative plasmid for the mating assays was the F derived plasmid, pOX38-Km.

Plasmid	Amino Acids of FinO	Color on MacLac	Miller Units	Transconj. / Donor	Mating Efficiency
pBC 4.0	186	white	0.0	0.00375	0.28%
pBC+1550*	186	white	0.0	-	-
pBC+972	186	white	-	-	-
pBC+877	186	white	0.0	0.018	1.3%
pBC+779	186	white	0.0	0.021	1.6%
pBC+484	161	red	969.5	2.50	185%
pBC+181	60	red	608.5	-	-
pBC+30	10	red	1038.2	-	-
pBC+18	6	red	1030.6	-	-
pBC-751	0	red	870.6	-	-
pBC-1050*	0	red	-	-	-
pBC-1650*	0	red	-	-	-
pBC-KS+	0	red	1133.3	1.35	100%

β-galactosidase assays were done by the method of Miller (1972) to measure the degree to which fertility inhibition had been affected by the various deletions of the *finO* region. Cultures were grown from single colonies of SE5000/pMCJ184/pBC 4.0 (or derivative). The culture was permeablized by toluene and then assayed for ONPG hydrolysis by β-galactosidase. β-galactosidase activities are described in Miller Units (M.U.) and are summarized in Table 3.1. Clones which have deletions extending into the *finO* coding region had considerable β-galactosidase activity, ranging from 609 to 1038 M.U. (pBC+181 and pBC+30 respectively). The negative control pBC-KS+ had a comparable activity of 1133 M.U. pBC+484, which is missing only 13% of the 3' coding region of *finO* had an equally high β-galactosidase activity (970 M.U.). Clones which encoded *finO* and *orf86* (pBC 4.0 and pBC+1550$^{\bullet}$) displayed no β-galactosidase activity, suggesting that the *finO* product was being expressed strongly. The evidence suggests that the *finO* gene product alone is able to confer full repression of *traJ::lacZ* translation, and that *orf86* does not play a regulatory role in *finO* expression since cells containing pBC+877 and pBC+779 displayed complete repression of β-galactosidase activity.

III. Mating Inhibition by R6.5 *finO* Deletion Derivatives

The previous tests assayed for inhibition of TraJ mRNA translation initiation as mediated by FinP RNA, which was in turn protected by FinO protein. It is also important to test the degree of repression by the *finO* constructs on the larger physiological scale: repression of the mating event. To test for mating efficiency, pBC plasmids were transformed into the *E. coli* strain MC4100/pOX38-Km. Because pOX38-Km carries the entire *tra* region except the *finO* gene, it is derepressed for transfer. The addition of *finO* in *trans* (i.e. on pBC plasmids) should result in the repression of transfer by pOX38-Km. The recipient used was the Spc^{R} strain ED24. The mating efficiency is based on the number of transconjugants produced per donor organism as compared to the strain MC4100/pOX38-Km/pBC-KS+ (see Table 3.1). The data for the clones are

consistent with the trends seen in the β-galactosidase assays, indicating that efficient transfer of pOX38-Km does not occur when *finO* is present; 0.3% transfer by pBC 4.0, 1.3% by pBC+877, 1.6% by pBC+779. The absence of *finO* (pBC-KS+) or a minimal truncation of the *finO* coding region (pBC+484) resulted in the derepression of transfer, allowing for a transfer efficiency that exceeds 100%.

Discussion

A mating efficiency of 0.3% was observed when the entire *Pst*I fragment from R6-5 was present, indicating that complete repression of transfer is imposed by FinO when supplied in *trans*. Deletions of the downstream sequences did give a minor increase in mating efficiency, 1.3% for pBC+877 and 1.6% for pBC+779. However these increases in mating efficiency are less than one log greater than that seen in the pBC 4.0 clone, suggesting that the effect is not physiologically relevant. From this it can be concluded that the open reading frame, *orf86*, and other downstream sequences do not play a significant role in the regulation of fertility inhibition. Indirect assays of FinP stability using the *lacZ* fusion construct, pMCJ184, show a similar trend. Clones containing the plasmids pBC 4.0, pBC+877 and pBC+779 displayed no β-galactosidase activity, showing that sufficient FinO protein was synthesized to protect FinP transcripts. Complete repression was observed in both the presence and absence of the sequences downstream of *finO*.

My results agree with those of previous studies which used the natural mutants R100-1 or F (pOX38-Km), which produce a 75 amino acid peptide or a 131 amino acid truncated peptide of FinO respectively. Both are deregulated for transfer (Sugino and Hirota, 1962; Cheah and Skurray, 1986). Deletions constructed in this study, which extend into the *finO* coding region also caused deregulation. The plasmid pBC+484, which produces a 161 amino acid peptide

of FinO displayed a loss of fertility inhibition, with a mating efficiency of 185%. Co-residency of pBC+484, with the *lacZ* fusion construct resulted in an increase of β-galactosidase activity to ~970 M.U. When assayed for β-galactosidase activity, plasmid pBC+181, which encodes a 60 amino acid peptide showed an activity of ~610 M.U. While this activity is lower than the clones carrying pBC-KS+ and pBC+484, the activity is still a great deal higher than the repressed clones (pBC 4.0, pBC+877). The mating data again support the conclusion that the downstream sequences are not necessary for fertility inhibition, and that only the *finO* open reading frame is required. While this evidence does not directly refute the Dempsey "FinO as an antisense RNA" model, it does support the classical "FinO as a protein" model. These data indicate that even a minor deletion of the open reading frame completely relieves repression. Furthermore, if an antisense *finO* transcript is synthesized from the opposite strand at the 5' end of the accepted *finO* open reading frame, then the observation of deregulation of transfer by the plasmid pBC+484 cannot be explained. If a second, 3' transcript (designated FinX in this thesis, Chapter 1) is necessary (Dempsey 1987), then it must be transcribed from a promoter within the bounds of the accepted *finO* open reading frame. Data obtained from *in vitro* work by van Biesen and Frost (1994) and this thesis (Chapter 4) argue further against the "FinO as an antisense RNA" model.

Chapter 4. Characterization of the Functional Domains of FinO, the Fertility Inhibition Protein of F-like Plasmids.

Introduction

Fertility inhibition in the F-like plasmids depends upon a two component negative regulator which represses expression from the *tra* operon. The first component of the fertility inhibition system, FinP, is a 78 base RNA (Mullineaux and Willetts, 1985) which is antisense to the *traJ* gene. The structure of FinP has been determined to consist of a short, 4 base 5' leader region, followed by two stem-loop structures, which are in turn separated by a 4 base single-stranded spacer. The 3' end of the RNA is unstructured, and terminates in a stretch of Us (van Biesen *et al.*, 1993). Because FinP RNA is antisense to *traJ*, it is able to base pair, and to potentially form a full-length duplex with TraJ mRNA (Mullineaux and Willetts, 1985; Finlay *et al.*, 1986; van Biesen *et al.*, 1993).

The second component of the fertility inhibition system is the 21.2 kDa FinO protein, which protects FinP from degradation, increasing the FinP half life from ~3 minutes to >40 minutes *in vivo* (Lee *et al.*, 1992). FinO has three identified functions: RNA binding, RNA protection, and mediation of duplex formation (van Biesen and Frost, 1994). The purpose of this study is to identify the regions or domains of the FinO protein that are responsible for each of these functions. It is also possible that FinO operates as a multimeric protein, and so a region or domain of FinO may be important for protein-protein interactions. This last possibility was not tested in these experiments.

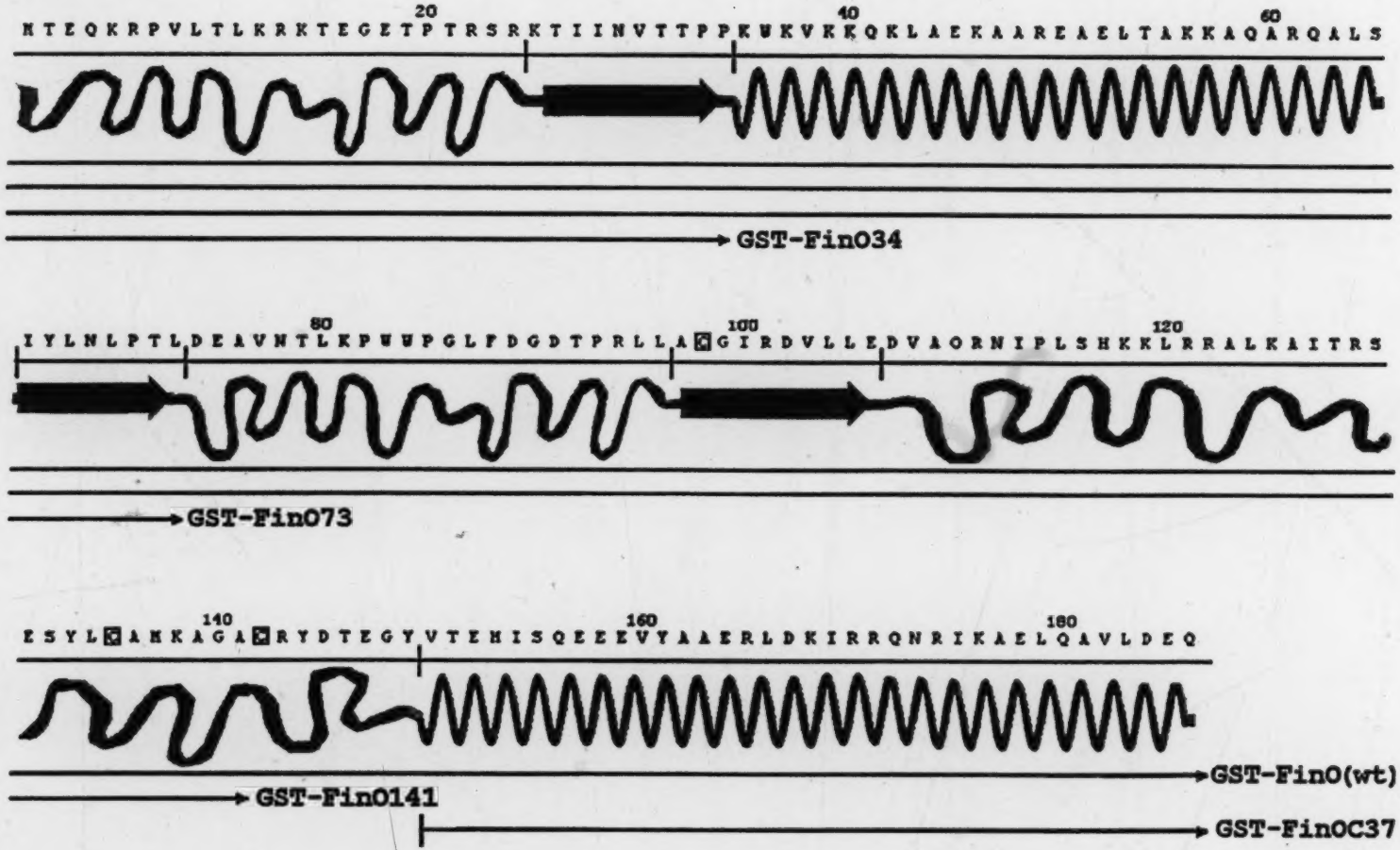
Results

I. Creation of Truncated GST-FinO Constructs

The amino acid sequence of FinO was analyzed by PLOTSTRUCTURE (GCG-Wisconsin package). Two α-helical domains were predicted by the program and the locations of these α-helical domains are noted in Figure 4.1. A domain of

Figure 4.1 The predicted secondary structure of the FinO protein. Regions that are predicted to fold as β-sheets are indicated by the solid arrow, while α-helical domains are indicated by the unbroken coil. Regions which have no predicted structure are indicated by the bent line. Each of the truncated proteins used in this study are depicted by long thin arrows below the schematic. All of the proteins used in this study were isolated as GST-fusions with the GST moiety located at the N-termini of the proteins. The C-terminus of each protein is indicated by a small arrowhead. Cysteine residues are boxed C in the sequence line.

Figure 4.1


20
40
60
80
100
120
140
160
180
GST-Fin034
GST-Fin073
GST-Fin0141
GST-Fin0(wt)
GST-Fin0C37

undefined structure and a short β-sheet domain are predicted to precede the first N-terminal α-helix. The region between the helices contains two predicted β-sheets and a region with no predicted structure. The C-terminal α-helical domain extends to the end of the protein. Many RNA binding proteins utilize basic amino acids, located in α-helices, to bind their RNA targets. The most notable example is the dimeric protein, Rom, which binds the target RNA I:RNA II complex via a number of very basic residues, which are aligned along one face of the 1/1' α-helical domain (Predki *et al.*, 1995). The putative N-terminal α-helix in FinO is very basic, with predicted pI of 11.2 but the C-terminal α-helix, with a predicted pI of 5.1 is moderately acidic. The region preceding the first α-helix (residues 1 through 34) is predicted to be very basic, with a pI of 11.7 (ISOELECTRIC, GCG-Wisconsin package). Prior to the experiments it seemed that the N-terminal α-helix and the region preceding it would be the most likely candidate for the RNA binding site(s). It was also possible that the C-terminal α-helix could play the role of an "electrostatic rudder" as was predicted for the 2/2' helix of Rom protein (Predki *et al.*, 1995). Several mutants were designed on the basis of the computer protein-folding results, in order to define the minimal region(s) required for RNA binding by the FinO protein.

The smallest of the fusion proteins produced in these experiments is synthesized from the plasmid pGEX-F034. The peptide corresponds to the first 34 amino acids of wild-type FinO, which does not include the N-terminal α-helix (see Figure 4.1). The peptide produced from pGEX-FO73 corresponds to the first 73 residues of wild-type FinO, which includes the predicted N-terminal α-helical domain. The three cysteine residues located in the region between the α-helices may be of some importance for protein folding (*ie.* disulfide bonds). To test for the function of the intervening peptide sequence, and to test for the possible role of the cysteine at position 142 (C142), the next peptide truncation was made at residue 141, approximately 8 residues before the C-terminal α-helical domain. The resulting peptide, GST-FinO141, is encoded by the plasmid pGEX-FO141.

The C-terminal α-helix was also tested for RNA binding activity in the absence of the other putative domains. It was hoped that the product, GST-FOC37, which corresponds to the last 37 residues of the FinO protein would fold in isolation as it does in the wild-type protein. As such, positive results for binding or protection (or duplex formation) would be meaningful, but any negative results could only be considered indeterminate because protein folding of this domain was not examined. However, due to the relative acidity of the C-terminal region, it seems unlikely that this region would be responsible for RNA binding activities.

The plasmid constructs were transformed into *E. coli* DH5α. Cultures were grown from single colonies in TSB, and expression of the GST fusion proteins was induced by IPTG. Figure 4.2 shows both induced and uninduced clones, which were resolved by electrophoresis on an 8% SDS-Polyacrylamide gel. The stain employed was Coomassie brilliant blue R-250.

II. Conjugation is not Inhibited by Truncated GST-FinO Products

The pGEX constructs were tested for their ability to inhibit transfer events by the deregulated F derivative, pOX38-Km. The pGEX plasmids were transformed into the *E. coli* strain MC4100 containing pOX38-Km. The previously described *finO*$^+$ plasmid pSnO104 (Lee *et al.*, 1992), and the immobile parental vector pACYC184 (Chang and Cohen, 1978) were employed as positive and negative controls respectively. The efficiency of mating was determined, and is equal to the number of transformants produced per donor cell after a 30 minute mating event. The efficiency was then expressed as the % transfer relative to the negative control containing pOX38-Km/pACYC184 (Table 4.1). The greater the degree of fertility inhibition, the lower the % transfer relative to the negative control. In these assays the control vectors pACYC184 and pGEX-2T did not inhibit the transfer of pOX38-Km to any significant degree. pSnO104 and pGEX-FO2 both inhibited transfer relative to the negative controls to a mating efficiency of <1%. This is in good agreement with previously published results for pGEX-

Figure 4.2 10% SDS-polyacrylamide gel of GST-fusion proteins prepared from *E. coli* DH5α. Lane 1 contains high molecular weight markers (BRL). From top to bottom the bands correspond to molecular weights of 103, 67.5, 44.5, 29.1, 18.9 and 15.7 kDa. Lane 2 contains the whole cell protein preparation from uninduced DH5α/pGEX-FO2. Lanes 4, 6, 8, 10 and 12 contain whole cell protein from uninduced DH5α harboring plasmids pGEX-FO141, pGEX-FO73, pGEX-FO34, pGEX-FOC37 and pGEX-2T respectively. Lanes 3, 5, 7, 9, 11 and 13 contain protein from IPTG induced DH5α harboring plasmids pGEX-FO2, pGEX-FO141, pGEX-FO73, pGEX-FO34, pGEX-FOC37 and pGEX-2T respectively. The gel was stained with Coomassie Brilliant Blue dye R-250.

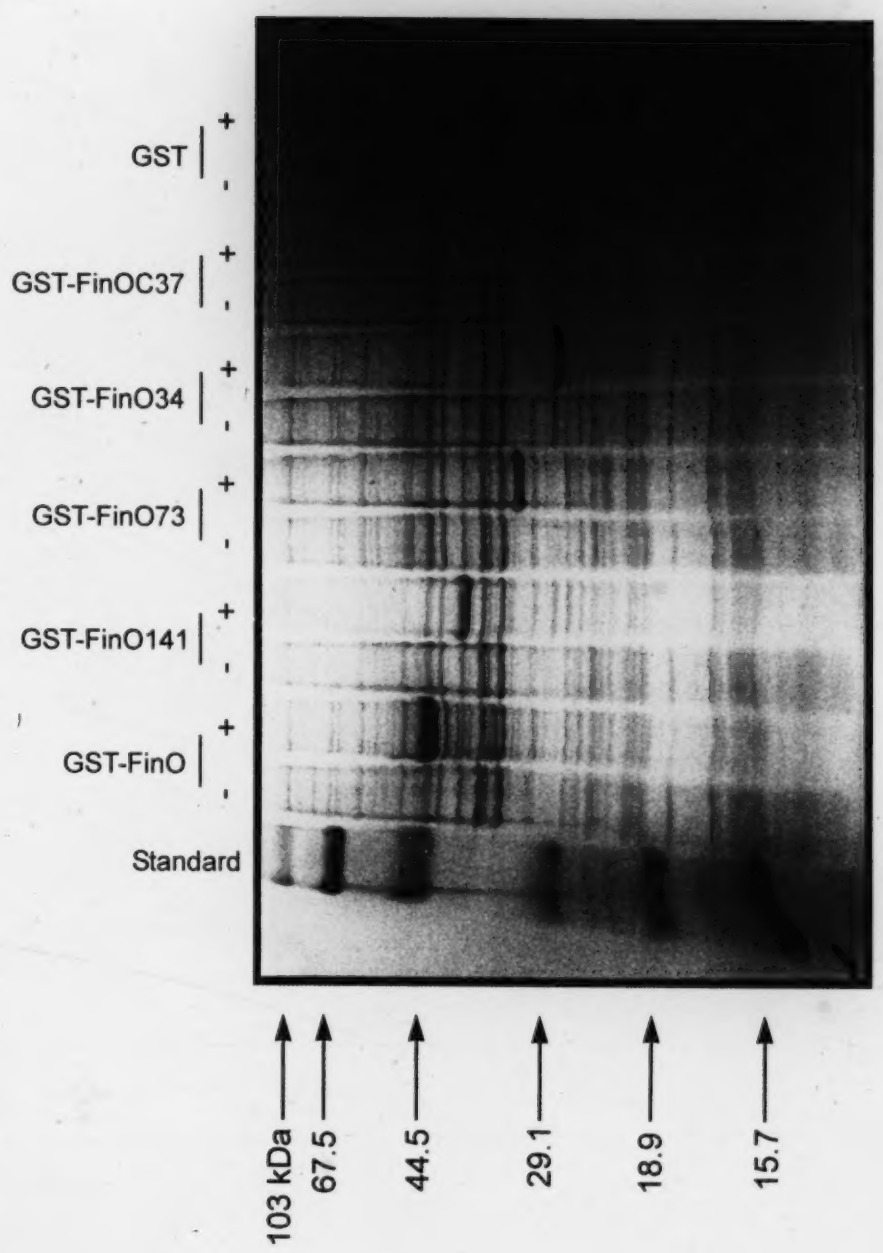

GST
+
-
GST-FinOC37
+
-
GST-FinO34
+
-
GST-FinO73
+
-
GST-FinO141
+
-
GST-FinO
+
-
Standard
103 kDa
67.5
44.5
29.1
18.9
15.7

Figure 4.2

Table 4.1 Mating efficiencies of plasmid pOX38-Km with co-resident plasmids. [a]pACYC184 and pGEX-2T do not encode a *finO* gene. [b]pSnO104 is a *finO*+ construct based on the vector pACYC184. [c]pGEX-FO2 and pGEX-FO141 encode the entire *finO* gene and a 24% C-terminal deletion of the *finO* gene respectively. The *finO* fragments were cloned into the multiple cloning site of the expression vector pGEX-2T. [d]Matings were done from the donor strain, *E. coli* MC4100, into the recipient, ED24. The number of transconjugants per donor organism is indicated. [e]Additionally, the efficiency of mating relative to the vector control, MC4100/pOX38-Km/pACYC184 is reported.

Donor Organism (MC4100)	Transconjugants per Donor Organism[d]	Mating Efficiency relative to pACYC184[e]
pACYC184/pOX38-Km[a]	1.88	100%
[b]pSnO104/pOX38-Km	.0013	0.07%
pGEX-2T/ pOX38-Km[a]	1.72	91.3%
pGEX-FO2/ pOX38-Km[c]	.0074	0.39%
pGEX-FO141/ pOX38-Km[c]	2.02	107.3%

FO2, where mating efficiencies dropped to 3-3.5% of wild-type (i.e. *finO*$^-$) levels (van Biesen and Frost, 1994). The mutant construct pGEX-FO141 gave a mating efficiency of 107.3%, indicating that all repression is lost. In previous work on R6-5 *finO*, an even smaller truncation of only 13% of the C-terminal coding region was sufficient to completely deregulate fertility inhibition (Chapter 2). The mutants pGEX-FO73, pGEX-FO34, and pGEX-C37 were not tested for their ability to inhibit transfer.

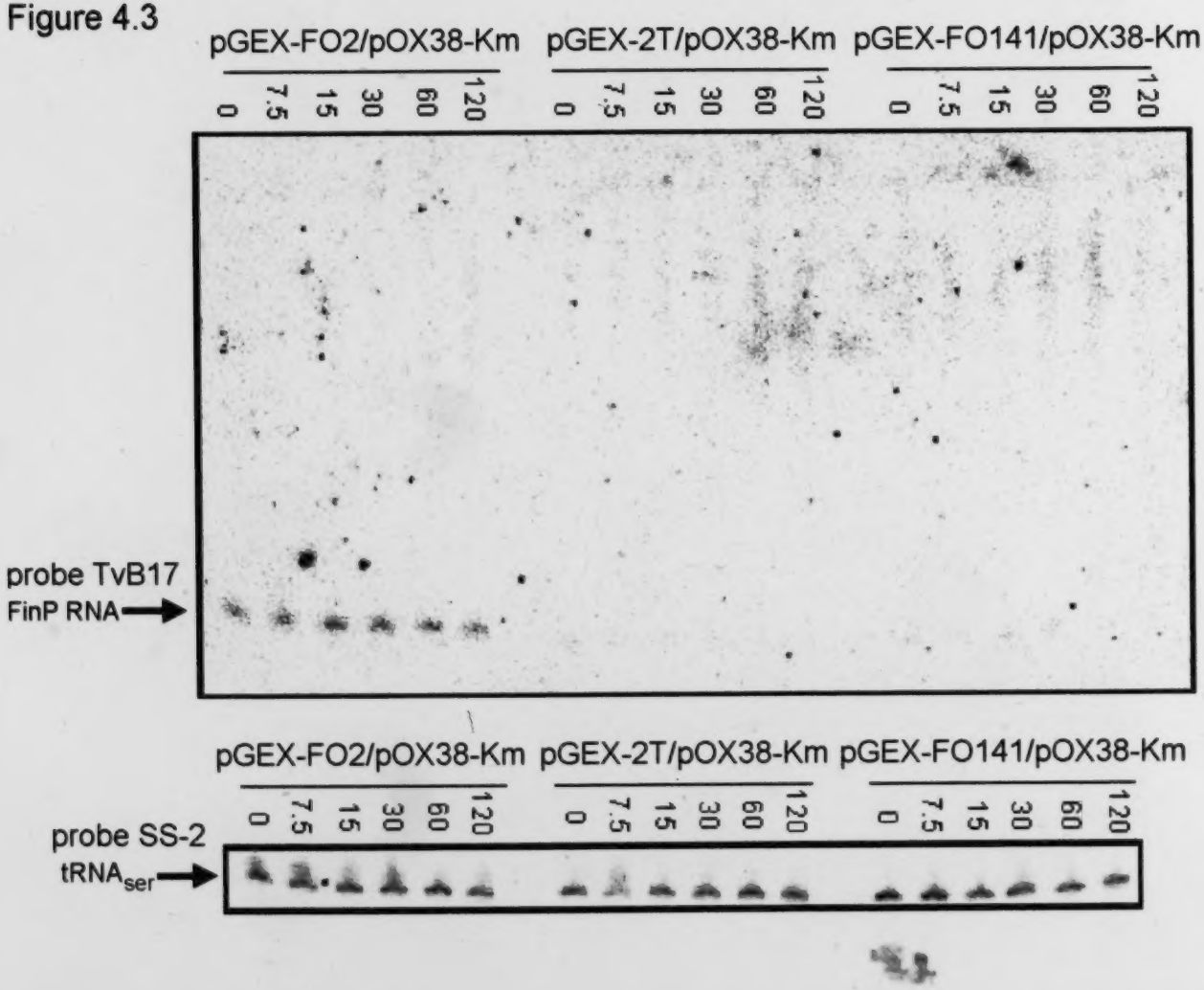
III. GST-FinO Derivatives do not Stabilize FinP RNA *In vivo*

GST-FinO protein, when expressed from the plasmid pGEX-FO2 was able to repress transfer of pOX38-Km, while the derivative GST-FinO141 was not able to repress transfer. It remained uncertain whether the binding function, the protection function or the annealing function had been disrupted in GST-FinO141, allowing for the observed increased rate of transfer. In this assay FinP transcript concentrations from pOX38-Km were determined for each of the GST-FinO derived fusions. Cells were treated with rifampicin and samples were taken at the time points indicated. Whole cell RNA was isolated, and 50 µg were loaded to an 8% denaturing polyacrylamide gel. RNA was transferred from the gel to nylon membranes, probed by a FinP-specific, [^{32}P]-labeled primer and exposed to a phosphor screen overnight. Membranes were then stripped and probed by a $tRNA_{ser}$ specific, [^{32}P]-labeled primer as an internal standard. Northern blots of cells containing plasmid pGEX-FO2, pGEX-2T, and pGEX-FO141 (co-resident with pOX38-Km) are shown in Figure 4.3.

GST-FinO provided by pGEX-FO2 stabilized FinP *in vivo*, allowing for a significant accumulation of the transcript prior to rifampicin treatment (see time point zero), and extended the half-life of the RNA to >60 minutes. The negative control pGEX-2T did not confer protection to FinP. Very little FinP was present at time zero, and RNA was not quantifiable after 60 minutes. It is interesting to note that FinP from the pOX38-Km/pGEX-2T clones migrates at a slightly

Figure 4.3 Northern hybridization analysis. Whole cell RNA preparations from *E. coli* SE5000 containing the plasmids pOX38-Km and pGEX-FO2 derivatives. Cultures were treated with rifampicin to stop transcription and samples were taken at time 0, 7.5, 15, 30, 60, and 120 minutes. Whole cell RNA was prepared, separated on a denaturing PAG and transferred to a nylon membrane. The membrane was probed with a FinP RNA-specific probe, TvB17. An arrow indicates the FinP band where TvB17 has hybridized. Note that FinP from pGEX-FO2-containing cells migrates slightly slower than the FinP from the pGEX-2T- and pGEX-FO141-containing cells. The membranes were stripped of TvB17 and reprobed with the $tRNA_{ser}$ specific probe, SS-2. The relevant portion of the SS-2 blot is aligned below, and was used to accurately assess the total cellular RNA loaded in each lane. FinP is ~80 nucleotides long, and tRNAser is 90 nucleotides in length (Steege, 1983).

Figure 4.3


pGEX-FO2/pOX38-Km
pGEX-2T/pOX38-Km
pGEX-FO141/pOX38-Km
0
7.5
15
30
60
120
probe TvB17
FinP RNA
probe SS-2
tRNA$_{ser}$

greater rate than FinP RNA from pOX38-Km/pGEX-FO2 clones. It is possible that some RNase cleavage of the RNA has occurred. Similarly pGEX-FO141 did not confer stability to FinP, indicating that the protein GST-FinO141 is unable to protect FinP from degradation *in vivo*. Cells containing pGEX-FO73, pGEX-FO34, and pGEX-FOC37 (co-resident with pOX38-Km) are shown in Figure 4.4. As with the clone pOX38-Km/pGEX-FO141 and the negative control, pOX38-Km/pGEX-2T, no stabilization of FinP was observed in these mutants. Lane 7 contains RNA from pOX38-Km/pGEX-FO2 from time zero as a positive control. As was observed with pGEX-2T and pGEX-FO141 in Figure 4.3, cells containing the mutant plasmids pGEX-FO73, pGEX-FO34 and pGEX-FOC37 produced low levels of a truncated FinP product which were detectable at early time points.

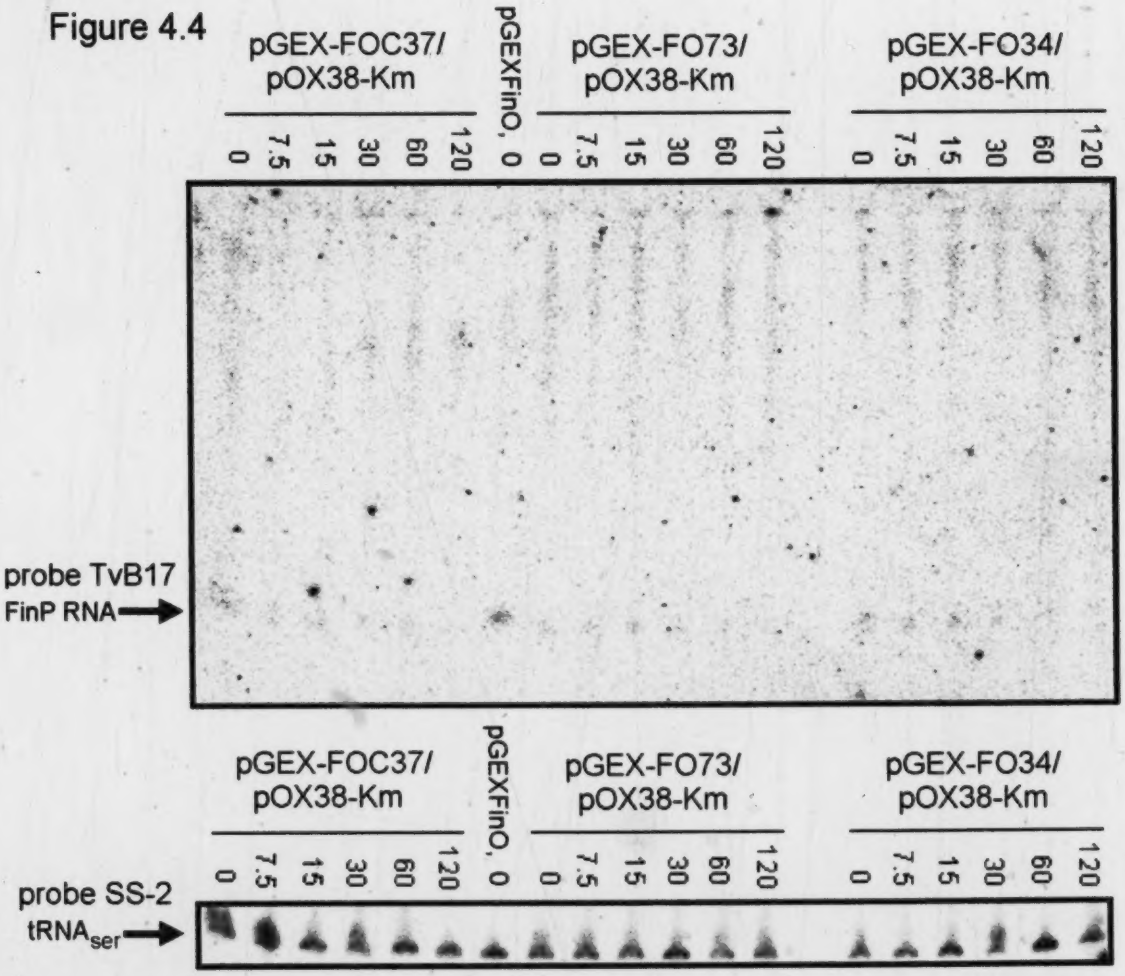
IV. The N-terminal α-helical Domain is Required for FinP Binding

All peptides including GST-FinO and the negative control, GST, were synthesized and purified by a procedure based on the methods of Frangioni and Neel (1993) and of van Biesen and Frost (1994). The GST moieties were not removed from purified GST-fusion peptides for reasons outlined by van Biesen and Frost (1994). In some protein preparations low levels of protein degradation were observed (data not shown). The degradation product in all protein samples migrated at the same rate as purified GST, thus the degradation products were not considered major contaminants. Negative controls included purified GST to ensured that the observed RNA binding was due to the activity of the fusion proteins only. Purified proteins were stored at –20°C until needed.

[^{32}P] UTP-labeled FinP transcripts were synthesized from the plasmid pLJ5-13. A truncated, [^{32}P] UTP-labeled TraJ mRNA was synthesized from PCR-generated DNA templates. It should be noted that the truncated mRNA product is identical to the product previously described as TraJ211 (van Biesen *et al.*, 1993; van Biesen and Frost, 1994), and that it has been renamed $TraJ_{184}$ because the correct length of the transcript is 184 bases. Varying concentrations of the GST-fusion

Figure 4.4 Northern hybridization analysis. Whole cell RNA preparations from *E. coli* SE5000 containing the plasmids pOX38-Km and pGEX-FO2 derivatives. Cultures were treated with rifampicin to stop transcription and samples were taken at time 0, 7.5, 15, 30, 60, and 120 minutes. Whole cell RNA was prepared, separated on a denaturing PAG and transferred to a nylon membrane. The membrane was probed with a FinP RNA-specific probe, TvB17. An arrow indicates the FinP band where TvB17 has hybridized. Note that FinP from pGEX-FO2-containing cells (lane 7) migrates slightly slower than the FinP from the pGEX-FOC37, pGEX-FO73 and pGEX-FO34-containing cells. The membranes were stripped of TvB17 and reprobed with the $tRNA_{ser}$ specific probe, SS-2. The relevant portion of the SS-2 blot is aligned below.

Figure 4.4


pGEX-FOC37/
pOX38-Km
pGEXFinO, 0
pGEX-FO73/
pOX38-Km
pGEX-FO34/
pOX38-Km
0
7.5
15
30
60
120
probe TvB17
FinP RNA
probe SS-2
tRNA$_{ser}$

proteins were incubated with the [^{32}P]-labeled FinP RNA at room temperature. The mixtures were then loaded to a non-denaturing polyacrylamide gel. The protein bound RNA was differentiated from the unbound RNA on the basis of its reduced rate of migration (mobility shift). The association constant, K_a, was calculated from the protein concentration which caused a mobility shift of half the labeled RNA. Association constants were determined for both the FinP and $TraJ_{184}$ transcripts. The GST-FinO K_a for FinP as reported previously by van Biesen and Frost (1994) is ~4 x 10^5 M^{-1}. In the FinP binding experiments reported here, GST-FinO caused a 50% shift of the target RNA at a concentration of 1.75 x 10^{-8} M (Figure 4.5), which corresponds to a K_a for FinP of 5.6 x 10^7 M^{-1} (Standard deviation of 1.8 x 10^6).

The FinP RNA mobility shift profile for GST-FinO141 can be seen in Figure 4.6. Significant aggregation of the protein occurred, which is not observed in the GST-FinO mobility shift experiments until very high concentrations of protein are used. The aggregates are too large to migrate through the polyacrylamide matrix, so remain in the wells. The calculated K_a for GST-FinO141 binding to FinP RNA is 3.5 x 10^7 M^{-1} (SD 7.8 x 10^5). When determining the K_a of GST-FinO73 some dissociation of the protein was observed. As a result the reported value of 2.2 x 10^7 M^{-1} (SD 9.4 x 10^6) for FinP binding is considered approximate. GST-FinO73 did not shift the FinP RNA to a discrete band as observed in both the GST-FinO and GST-FinO141 gels, but formed aggregates in the wells (Figure 4.7). However, the binding of FinP by the aggregated protein remained specific as shown by tRNA binding competitions (see below). In Figure 4.8 we can see that GST-FinO34 was supplied in vast molar excess (~11 000 fold excess, lane 1) but did not cause a detectable mobility shift of FinP RNA. Likewise GST-FinOC37 was supplied in molar excess (~18 000 fold excess, lane 1, Figure 4.9) but did not cause a detectable mobility shift of FinP RNA. Finally, it can be seen that 3.3 x 10^{-6} M GST (a molar ratio of ~7000:1 GST to labeled RNA) is unable to cause any notable mobility shift of the FinP transcripts (Figure 4.10).

Figure 4.5 Association of GST-FinO with [^{32}P]-labeled FinP or TraJ184 RNA, as indicated. Varying amounts of GST-FinO protein (pmole) were incubated with 5 x 10^{-10} M FinP RNA (or TraJ184 RNA) at room temperature for 30 minutes. Samples were resolved on an 8% non-denaturing polyacrylamide gel and scanned by a phosphorimager. K_a values were calculated from this gel and others like it.

Figure 4.6 Association of GST-FinO141 with [^{32}P]-labeled FinP or TraJ184 RNA, as indicated. Varying amounts of GST-FinO141 protein (pmole) were incubated with 5 x 10^{-10} M FinP RNA (or TraJ184 RNA) at room temperature for 30 minutes. Samples were resolved on an 8% non-denaturing polyacrylamide gel and scanned by a phosphorimager.

Figure 4.5

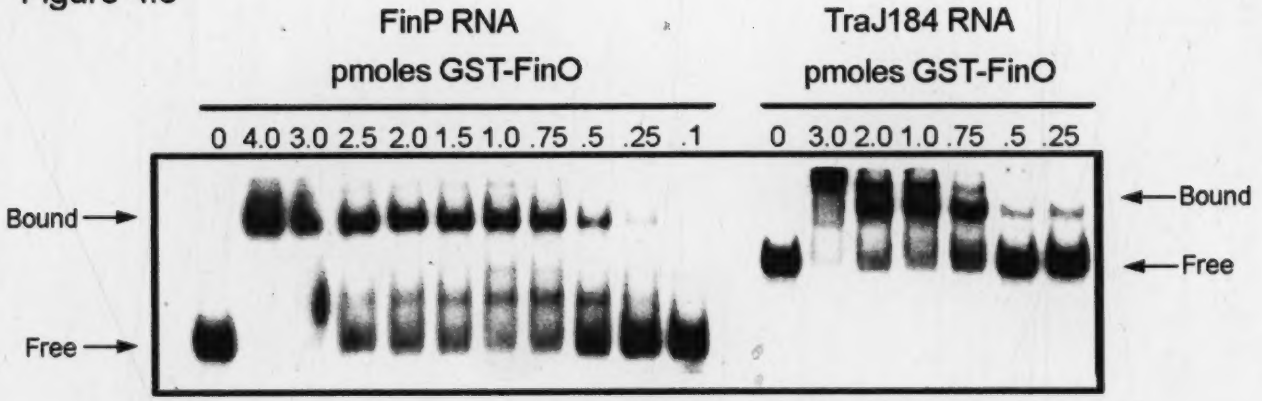


Figure 4.6

Figure 4.7 Association of GST-FinO73 with [^{32}P]-labeled FinP or TraJ184 RNA, as indicated. Varying amounts of GST-FinO73 protein (pmoles) were incubated with 5 x 10^{-10} M FinP RNA (or TraJ184 RNA) at room temperature for 30 minutes. Samples were resolved on an 8% non-denaturing polyacrylamide gel and scanned by a phosphorimager.

Figure 4.7

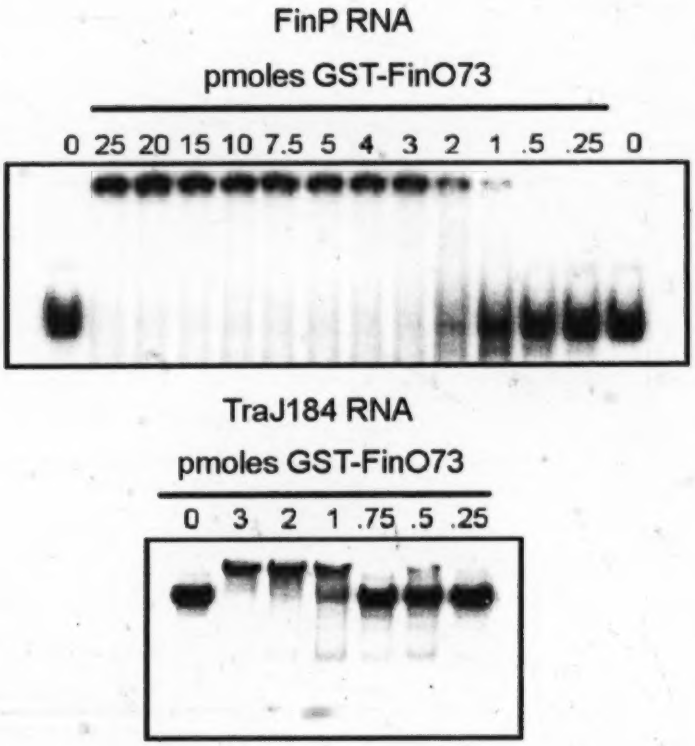

FinP RNA
pmoles GST-FinO73
0 25 20 15 10 7.5 5 4 3 2 1 .5 .25 0
TraJ184 RNA
pmoles GST-FinO73
0 3 2 1 .75 .5 .25

Figure 4.8 Co-incubation of GST-FinO34 with 5 x 10^{-10}M [^{32}P]-labeled FinP RNA for 30 minutes at room temperature. Samples were resolved on an 8% non-denaturing polyacrylamide gel and scanned by a phosphorimager. Ka values could not be calculated because no mobility shift was observed.

Figure 4.9 Co-incubation of GST-FinOC37 with 5 x 10^{-10}M [^{32}P]-labeled FinP RNA for 30 minutes at room temperature. Samples were resolved on an 8% non-denaturing polyacrylamide gel and scanned by a phosphorimager. Ka values could not be calculated because no mobility shift was observed.

Figure 4.8

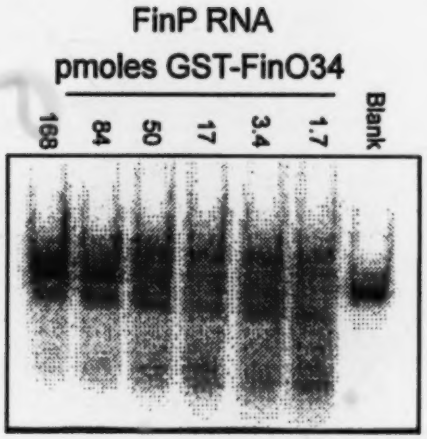


Figure 4.9

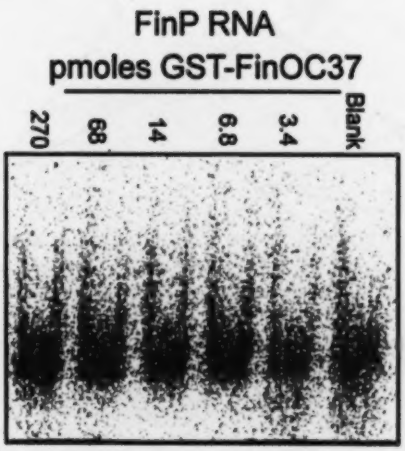
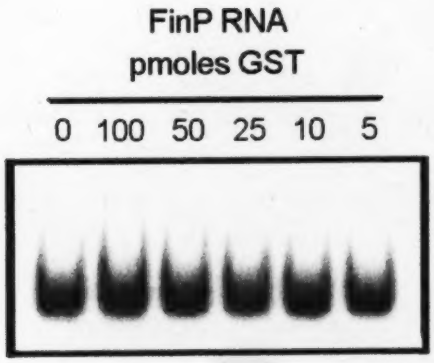


Figure 4.10 Co-incubation of glutathione-S-transferase (GST) with 5 x 10^{-10}M-[^{32}P] labeled FinP RNA for 30 minutes at room temperature. Samples were resolved on an 8% non-denaturing polyacrylamide gel and scanned by a phosphorimager.

Figure 4.10



The association constants for protein binding to the $TraJ_{184}$ transcripts are similar to the K_a values reported for FinP RNA. GST-FinO, GST-FinO141, and GST-FinO73 had $TraJ_{184}$ K_a values of 5.0×10^7, 5.0×10^7, and 3.5×10^7 M^{-1} respectively. The $TraJ_{184}$ K_a values reported are from single determinations. The abilities of GST-FinO34 and GST-FinOC37 to bind $TraJ_{184}$ were not tested.

V. FinP RNA Binding by GST-FinO Derivatives is Specific

Previous work with GST-FinO showed that the binding activity of the protein is moderately specific for FinP RNA (van Biesen and Frost, 1994) In the experiments presented here the proteins GST-FinO, GST-FinO141, and GST-FinO73 were examined for their specificity of binding to the FinP target. Specificity is defined as the preferential binding by the protein to FinP when co-incubated with excess quantities of *E. coli* derived tRNA. The protein/RNA reaction mixtures were incubated at RT for 30 minutes, then resolved by non-denaturing PAGE. The amount of excess tRNA required to efficiently compete for protein binding sites versus the labeled FinP correlates with the degree of specificity of the protein. As a positive control [^{3}H]-labeled FinP (termed "unlabeled") was added in excess of [^{32}P]-labeled FinP to demonstrate that a competition for protein binding sites does occur. The protein concentration which caused a 100% gel shift of labeled FinP was determined for each protein (section IV), and these concentrations were used in the respective specificity assays.

The GST-FinO specificity assay is shown in Figure 4.11. The % protein-bound [^{32}P]-FinP (shifted) and the % protein-free [^{32}P]-FinP (free) were calculated for each lane. From the control lanes, 50% of the labeled FinP RNA was free when a ratio of ~5:1 unlabeled to labeled FinP was added to GST-FinO protein. In the same experiment tRNA needed to exceed labeled FinP RNA by 85-fold to free 50% of the FinP from the RNA:protein complex. This confirms that GST-FinO does have a moderate specificity for FinP RNA. Radiographs of the specificity assays for the proteins GST-FinO73 and GST-FinO141 are shown in Figure 4.12

Figure 4.11 Specificity of binding by the GST-FinO protein. GST-FinO was co-incubated with 5 x 10^{-10} M [^{32}P]-labeled FinP and varying amounts of the specific competitor, [^{3}H]-FinP, or the non-specific competitor, unlabeled tRNA. The labeled substrate and the competitor transcripts were mixed prior to the addition of GST-FinO. The molar ratio of [^{32}P]-labeled FinP to competitor RNA is indicated at the top of each lane. The incubation conditions and gel electrophoresis methods are identical to those employed in the binding studies.

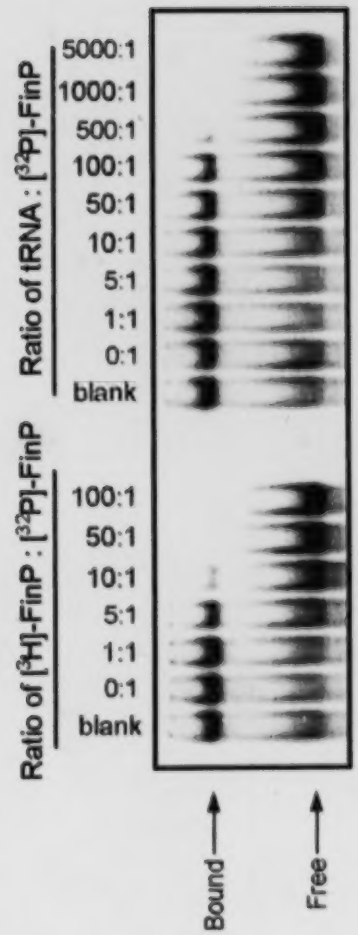
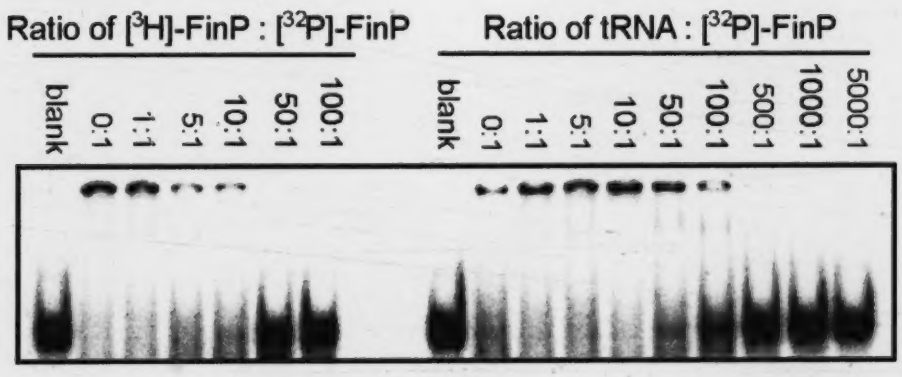

Ratio of [3H]-FinP : [32P]-FinP
100:1
50:1
10:1
5:1
1:1
0:1
blank
Ratio of tRNA : [32P]-FinP
5000:1
1000:1
500:1
100:1
50:1
10:1
5:1
1:1
0:1
blank
Bound
Free

Figure 4.11

Figure 4.12 Specificity of binding by the GST-FinO73 protein. GST-FinO73 was co-incubated with 5 x 10^{-10} M [^{32}P]-labeled FinP and varying amounts of the specific competitor, [^{3}H]-FinP, or the non-specific competitor, unlabeled tRNA. The labeled substrate and the competitor transcripts were mixed prior to the addition of GST-FinO73. The molar ratio of [^{32}P]-labeled FinP to competitor RNA is indicated at the top of each lane. The incubation conditions and gel electrophoresis methods are identical to those employed in the binding studies.

Figure 4.12


Ratio of [3H]-FinP : [32P]-FinP
blank
0:1
1:1
5:1
10:1
50:1
100:1
Ratio of tRNA : [32P]-FinP
blank
0:1
1:1
5:1
10:1
50:1
100:1
500:1
1000:1
5000:1

and Figure 4.13 respectively. 50% of the labeled RNA became free of GST-FinO73 protein in the control experiment at a ratio of 5:1 unlabeled to labeled FinP. A ratio of 100:1 tRNA to labeled FinP was needed to free 50% of the labeled RNA from GST-FinO73. As in the binding assays, GST-FinO73 formed aggregates which remained in the well. Despite aggregation, we can see in these experiments that the GST-FinO73 protein could still bind labeled FinP RNA with high affinity and with moderate specificity. Because GST-FinO141 is intermediate between GST-FinO73 and GST-FinO in size, it is expected that it will have a similar degree of specificity for the FinP target. A molar ratio of >100:1 tRNA to FinP is needed to free 50% of the labeled FinP. As such, GST-FinO141 follows the trend of specificity as seen with GST-FinO and GST-FinO73.

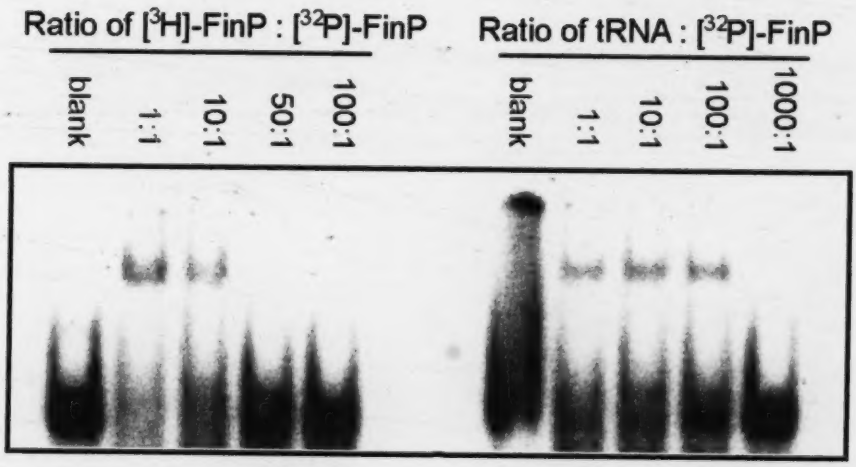
VI. Duplex Formation Rates are Improved by GST-FinO Derivatives

In addition to stabilizing FinP RNA *in vivo*, FinO functions to inhibit fertility by increasing the apparent rate of complex/duplex formation (k_{app}) between the antisense RNA, FinP, and the target, TraJ mRNA. FinO has been shown to increase the k_{app} between FinP and TraJ RNAs ~5-fold *in vitro* (van Biesen and Frost, 1994). The complex formation rate has not been determined *in vivo*, partially because the FinP/TraJ duplex is rapidly degraded by RNase III. In these studies, the objective was to confirm the previously published rates of complex formation and to determine the minimum peptide sequence required for the duplex promotion activity.

The rate of complex formation was determined by incubating $TraJ_{184}$ and FinP RNA at 37°C in a duplex formation buffer in the presence or absence of protein. During the 10 to 15 minute period of incubation, aliquots were removed and the RNA:RNA binding reaction was stopped by incubation in 1 x formamide solution at 4°C. The formamide treated aliquots were run on an 8% non-denaturing polyacrylamide gel. Because the FinP:$TraJ_{184}$ complexes are larger, they can be

Figure 4.13 Specificity of binding by the GST-FinO141 protein. GST-FinO141 was co-incubated with 5 x 10^{-10} M [^{32}P]-labeled FinP and varying amounts of the specific competitor, [^{3}H]-FinP, or the non-specific competitor, unlabeled tRNA. The labeled substrate and the competitor transcripts were mixed prior to the addition of GST-FinO141. The molar ratio of [^{32}P]-labeled FinP to competitor RNA is indicated at the top of each lane. The incubation conditions and gel electrophoresis methods are identical to those employed in the binding studies.

Figure 4.13


Ratio of [3H]-FinP : [32P]-FinP
blank
1:1
10:1
50:1
100:1
Ratio of tRNA : [32P]-FinP
blank
1:1
10:1
100:1
1000:1

differentiated from the smaller, more rapidly migrating uncomplexed RNAs on the polyacrylamide gels. The apparent second-order rate of complex formation (k_{app}) was calculated from the decrease in the % free FinP over time (see Materials and Methods for details). Figure 4.14 shows a time course of FinP:$TraJ_{184}$ complex formation in the blank control. The reaction mixture contained no protein, so the calculated k_{app} is the spontaneous rate of association between the sense and antisense transcripts. The negative control for complex formation, which contains the protein GST, is also shown in Figure 4.14. The same rate of association was observed in the blank and negative controls, indicating that the GST did not promote complex formation. The k_{app} for FinP and TraJ, as calculated from the blank and negative control experiments is 8.4 x 10^4 $M^{-1}s^{-1}$ (SD 3.1 x 10^4) which is close to the previously reported value of ~5 x 10^4 $M^{-1}s^{-1}$ (van Biesen and Frost, 1994). These values differ from the first published k_{app} for FinP and $TraJ_{184}$ (van Biesen *et al.*, 1993) due to differing buffer conditions (van Biesen and Frost, 1994).

All of the GST-fusion peptides were tested for their ability to promote duplex formation at a protein concentration of 1.7 x 10^{-7} M. This protein concentration is sufficient to shift 100% of the labeled FinP in binding assays (see section IV) for all of the GST-peptides tested. GST-FinO (Figure 4.15) increased the k_{app} of FinP/$TraJ_{184}$ to 3.1 x 10^5 $M^{-1}s^{-1}$ (SD 1.3 x 10^5), a 3.7 fold increase in complex formation relative to the control k_{app}. This value is somewhat lower, but comparable to the ~5-fold increase in k_{app} observed by van Biesen and Frost (1994). GST-FinO73 (Figure 4.16) displayed a similar k_{app} of 2.3 x 10^5 $M^{-1}s^{-1}$ (SD 8.5 x 10^3), a 2.7-fold greater k_{app} relative to the control experiments. The intermediate sized peptide, GST-FinO141 (Figure 4.16) showed the unexpectedly high rate of complex formation of 1.6 x 10^6 $M^{-1}s^{-1}$ (SD 2.5 x 10^5) which is 19-fold greater than the k_{app} observed in the controls, and is ~5-fold greater than the k_{app} seen in the GST-FinO experiments. The functional properties of each of the GST-fusion proteins are summarized in Table 4.2.

Figure 4.14 Duplex formation between [^{32}P]-labeled FinP and [^{3}H]-labeled TraJ184 transcripts. 5 x 10^{-10} M FinP was co-incubated at 37°C with 10-fold excess TraJ184 RNA. GST protein was added at time zero. No extra protein was added to the blank control. Samples were removed at the times indicated, and diluted in formamide loading buffer. Gel electrophoresis and quantitation methods are identical to those employed in the binding studies.

Figure 4.14

Glutatione-S-transferase
time in minutes
1 2 4 6 8 10
blank (no protein)
time in minutes
1 4 6 8 10
Duplex
Free

Figure 4.15 Duplex formation between [^{32}P]-labeled FinP and [^{3}H]-labeled TraJ184 transcripts. 5 x 10^{-10} M FinP was co-incubated at 37°C with 10-fold excess TraJ184 RNA. GST-FinO was added at time zero. Samples were removed at the times indicated, and diluted in formamide loading buffer. Gel electrophoresis and quantitation methods are identical to those employed in the binding studies.

Figure 4.15

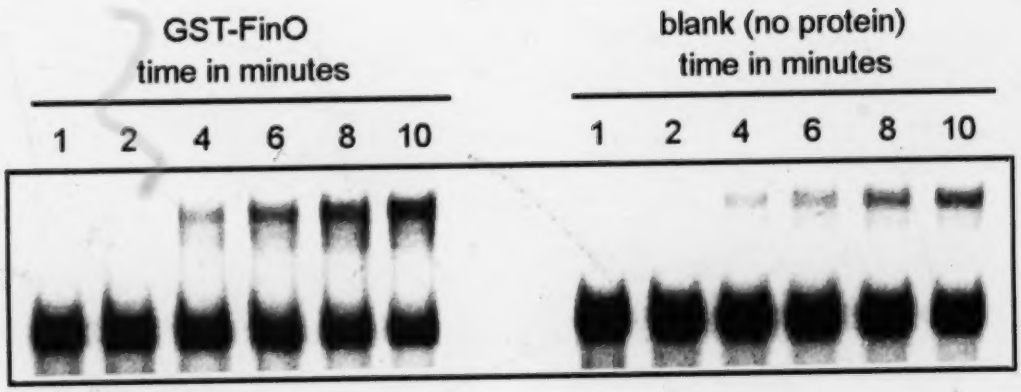

GST-FinO
time in minutes
1 2 4 6 8 10
blank (no protein)
time in minutes
1 2 4 6 8 10

Figure 4.16 Duplex formation between [^{32}P]-labeled FinP and [^{3}H]-labeled TraJ184 transcripts. 5 x 10^{-10} M FinP was co-incubated at 37°C with 10-fold excess TraJ184 RNA. GST-FinO73 or GST-FinO141 were added at time zero. Samples were removed at the times indicated, and diluted in formamide loading buffer. Gel electrophoresis and quantitation methods are identical to those employed in the binding studies.

Figure 4.16

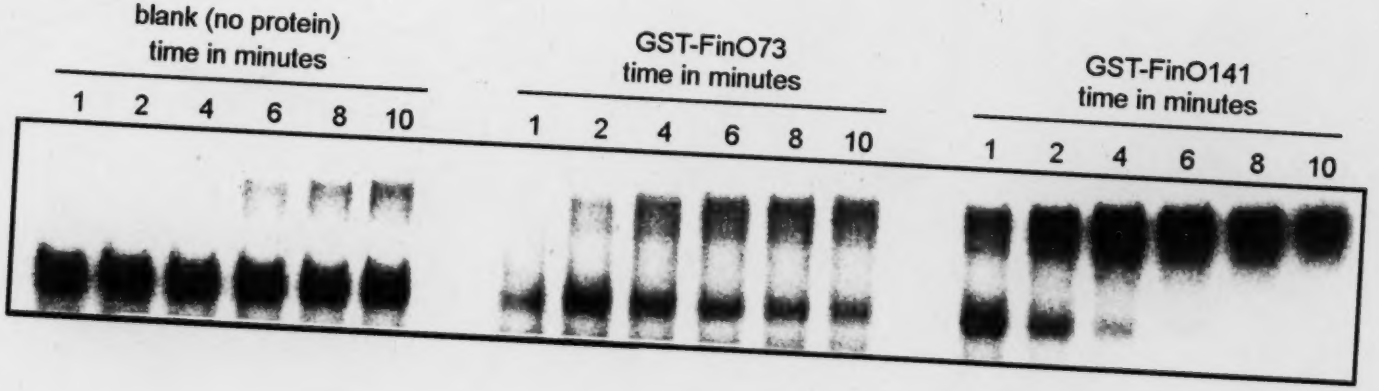

blank (no protein)
time in minutes
1 2 4 6 8 10
GST-FinO73
time in minutes
1 2 4 6 8 10
GST-FinO141
time in minutes
1 2 4 6 8 10

Table 4.2 Functional properties of GST-FinO and derivatives. [a]The association constant, K_a, of each protein for FinP RNA is indicated. Proteins which did not bind FinP are indicated by the (-) symbol. [b]The apparent rate of duplex formation, k_{app}, of each protein is indicated. Fusion proteins which were not tested are denoted by (n/d). FinP stability *in vivo* was determined by Northern analysis[c]. Clones which accumulated FinP due to protection from nucleases are indicated by the (+) sign, while clones which contained barely detectable levels of FinP are denoted by the (-) sign.

Fusion Protein	K_a (M^{-1})[a]	k_{app} ($M^{-1}s^{-1}$)[b]	FinP stability[c]
GST-FinO	5.6×10^7	3.1×10^5	+
GST-FinO141	3.5×10^7	1.6×10^6	-
GST-FinO73	~2.2×10^7	2.3×10^5	-
GST-FinO34	-	n/d	-
GST-FinOC37	-	n/d	-
GST	-	8.4×10^4	-

Discussion

In vitro work on the fusion protein GST-FinO by van Biesen and Frost (1994) described for the first time the molecular mechanism by which the *finO* gene product mediates fertility inhibition. FinO protein was shown to be an RNA binding protein that acts by binding and stabilizing the FinP transcript from degradation. The protein also acts to increase the rate of duplex formation between FinP RNA and the target mRNA, TraJ. In this study, a number of truncated GST-fusion proteins were produced using the commercial vector, pGEX-2T. The effects of the GST-FinO derivatives were studied *in vivo* and *in vitro*.

FinO was expressed as a fusion protein with the GST-moiety fused to the N-terminal end of the protein using the pGEX-based construct, pGEX-FO2. The transfer efficiency of pOX38-Km, a self-transmissible (*finO*$^-$) derivative of F, was assessed in the absence of *finO* using the vector control, pGEX-2T, or in the presence of the *finO* gene, as supplied by pGEX-FO2. It was previously reported that the efficiency of pOX38-Km transfer drops from 100% as seen in the vector control to ~3.5% when co-resident with pGEX-FO2 (van Biesen and Frost, 1994). In this report repression of transfer was also observed, though the efficiency of transfer dropped to <1% of the levels seen in the vector control. Previous work using exonuclease generated deletions of the *finO* coding region indicated that a truncation of the C-terminus by as little as 13% was sufficient to cause complete derepression of transfer of pOX38-Km (Chapter 3). This indicates that the C-terminal end is required for the FinO+ phenotype. pGEX-FinO141, which has 24% of the coding region deleted from the C-terminal end, was examined for the ability to repress the transfer of pOX38-Km. Not surprisingly, pGEX-FinO141 was unable to repress transfer of pOX38-Km when provided as a co-resident plasmid. The other pGEX-fusion constructs, which carry larger deletions than pGEX-FinO141were not examined for their ability to repress transfer.

The mating data indicate that the FinO mutants are incapable of repressing transfer. However the mechanism by which fertility control has been lost could not be addressed, due to the limitations of the experimental procedure. To examine whether the FinO mutants were able to stabilize FinP RNA *in vivo*, the cellular RNA concentrations were tested by the Northern blotting method. It was observed that GST-FinO, when provided in *trans*, stabilized the antisense RNA, and allowed for the accumulation of the transcript in the cells. Mutants which had any portion of the FinO C-terminus removed were unable to accumulate FinP at concentrations greater than the negative control, pGEX-2T. Furthermore no increase in the half-life of FinP was observed when mutants of FinO were supplied in *trans*, indicating that the N-terminal portion of the protein is critical for FinP protection *in vivo*.

The *in vitro* experiments assessed three aspects of the molecular functions of the FinO protein: binding ability, specificity of binding, and the ability to promote duplex formation. The truncation mutants were tested for their ability to bind FinP RNA, and the complementary transcript, TraJ mRNA. Of the mutants proteins tested, only those which displayed an RNA binding ability were further assayed for the specificity of binding, and for their effect on the rate of duplex formation.

In this study GST-FinO bound to FinP RNA with an association constant, K_a, of $5.6 \times 10^7\ M^{-1}$. This value is ~140-fold greater than the previously published value of $4.0 \times 10^5\ M^{-1}$ (van Biesen and Frost, 1994). The expected molar ratio of a binding protein to its target is 1:1, assuming the protein acts as a monomer. Prior work utilizing gel filtration chromatography suggested that wild type FinO, at least in the absence of target RNA, naturally exists in the monomer form (van Biesen and Frost, 1994). Because protein purification by the sonication and French press methods can cause significant damage to the proteins being purified, it is possible that the observed K_a values are artificially low. However, when comparing the isolation procedures for GST-FinO, it seems that the French press

method (this study) is preferable to the sonication method employed previously (van Biesen and Frost, 1994). Storage conditions are not suspected to cause a loss of activity since samples that had been stored at –20°C for as many as 30 days displayed the same binding rates as freshly prepared protein. Finally, binding by the fusion proteins may be sub-optimal due to the location of the fused GST moiety. As discussed below, the RNA binding domain is located at the N-terminus of the FinO protein, and this region is directly fused to the GST moiety. As a result, the GST peptide may sterically hinder RNA binding. Future studies on binding affinity should employ a C-terminal fusion of GST, or a histidine tag to reduce possible steric hindrance, and allow for relatively rapid and pure isolation of intact, active protein. It should be noted that several RNA binding proteins are able to multimerize *in vivo* (Zapp *et al.*, 1991; McCormick *et al.*, 1981; Friedman *et al.*, 1993). Some of these proteins are able to form multimers prior to RNA binding, while multimerization of others follows binding. It remains a possibility that FinO binds FinP as a multimer, or that the protein binds cooperatively, as is seen with Rev protein of HIV-1 (Daly *et al.*, 1993a). Some evidence for multimerization and possible cooperative binding of FinO to FinP/TraJ hybrids exists (van Biesen and Frost, 1994).

GST-FinO protein was shown by van Biesen and Frost (1994) to have a moderate specificity for FinP RNA. However, due to the low precision of the competition assay an accurate assessment of the degree of specificity of the protein was not possible. In this study GST-FinO was determined to bind FinP RNA with a ~17-fold preference versus the competitor, *E. coli* tRNA.

The truncated fusion proteins, GST-FinO141 and GST-FinO73 are both capable of binding FinP RNA with a K_a of 3.5 x 10^7 M^{-1} and ~2.2 x 10^7 M^{-1} respectively. GST-FinO141 formed both a discretely shifted RNA:protein complex and an RNA:protein aggregate complex. While it may be initially attractive to suggest that the RNA binding observed in the protein:RNA aggregates is non-specific, competition assays indicate that the aggregated mutant proteins have the same

specificity of binding as the wild type GST-FinO protein. The association constant of GST-FinO141 for FinP was determined to be 3.5 x 10^7 M^{-1}, which is 38% lower than that of wild type GST-FinO. This suggests that the C-terminal helical domain does not play a major role in RNA binding, but may play an important role in determining protein structure. Alternative conformations of the protein may be particularly prone to aggregate formation. Incubation of FinP RNA with GST-FinO73 resulted in the formation of RNA:protein aggregates in the wells. The competition assays indicate that the binding activity by the aggregated GST-FinO73 protein was no less specific than binding by wild type GST-FinO. The K_a of ~2.2 x 10^7 M^{-1} is only 39% of the wild type GST-FinO association constant. However, it should be noted that significant smearing of the labeled transcript was observed below the aggregates, suggesting that the RNA is dissociating from the protein complex at a rate greater than that observed with the other fusion proteins. While the central region of the protein does not play a direct role in binding the target RNA, it might act to stabilize the RNA:protein complex. This stabilization may occur in a manner similar to the Tat protein, which has a major, arginine-rich binding domain (Weeks *et al.*, 1990) and a secondary α-helical binding motif which is important for the stabilization of binding (Loret *et al.*, 1992). Overall, the unstructured central region of FinO (residues 74 – 141) is fairly basic in nature, and contains a run of very basic residues, HKKLRR, at positions 117 to 122. The central region of the protein should be further examined for a possible secondary RNA binding motif.

The association constants of the GST-fusion proteins for the $TraJ_{184}$ transcript were also tested. GST-FinO, GST-FinO141, and GST-FinO73 had K_a values of 5.0 x 10^7 M^{-1}, 5.0 x 10^7 M^{-1}, and 3.5 x 10^7 M^{-1} respectively. While these values differ somewhat from the K_a values obtained for FinP binding, it should be noted that the reported TraJ association constants were determined from a single experiment only. Nevertheless, these assays demonstrate clearly that two separate RNA binding domains for the FinP and TraJ transcripts do not exist. Rather, there is a single domain located at the N-terminal end of the protein, which is

capable of binding both the sense and antisense transcripts with approximately the same affinity.

No degree of mobility shift was observed when FinP was incubated with a vast molar excess of GST-FinO34, GST-FinOC37 or GST. From these results it is concluded that the first 73 amino acids contain the major binding domain of FinO. This minimal region allows efficient association with the FinP and TraJ targets, and binds these targets with the same specificity as the wild type protein. While the 34 N-terminal residues are not able to bind RNA themselves, a role in protein stabilization has not be ruled out. Further tests should be done with the first predicted α-helical domain (residues 35-73) fused directly to GST in the absence of the preceding residues. This will allow us to more fully determine the minimal sequence required for binding, and could help to elucidate the role, if any, of the extreme N-terminal end of the protein. The amino acid residues that are critical to RNA binding may be determined by the alanine-scanning mutagenesis method (Cunningham and Wells, 1989). In this method, residues within an α-helical domain are sequentially exchanged by site-directed mutagenesis to alanines. Because of the propensity of alanines to form α-helices, the integrity of the helix is unlikely to be disturbed. Residues which, when substituted by alanine, cause a significant decrease (or increase) in the association constant may directly contact the target RNA.

The last function examined was the ability of the GST-fusion proteins to promote duplex formation between the sense and antisense transcripts. The ability of FinP and TraJ to form duplexes was first tested in the presence of the GST protein, and separately in the complete absence of any protein. The rates (k_{app}) of both are reported to be 8.4×10^4 $M^{-1}s^{-1}$, indicating that GST alone is incapable of promoting duplex formation. In the previous communication by van Biesen and Frost (1994), the spontaneous rate of duplex formation in the presence of GST protein was reported to be slightly lower, at ~5×10^4 $M^{-1}s^{-1}$. In that same report, GST-FinO increased the rate of hybridization nearly 5-fold which corresponds to

a k_{app} of 2.4 x 10^5 $M^{-1}s^{-1}$. In this study GST-FinO is reported to have a k_{app} of 3.1 x 10^5 $M^{-1}s^{-1}$, which corresponds to an increase in the rate of duplex formation of 3.7-fold, as compared to the spontaneous rate of hybridization. The smallest mutant examined, GST-FinO73, promoted duplex formation by ~3-fold, with a k_{app} of 2.3 x 10^5 $M^{-1}s^{-1}$. From these results it is concluded that the minimal region which is capable of binding RNA is also able to promote duplex formation to a level comparable to that of the wild type protein. It is possible that conformational changes in the RNA structure occur during protein binding, as seen in the Tat/TAR system (Zacharias and Hagerman, 1995). Conformational changes could allow for the observed increase in the rate of RNA hybridization.

The duplex formation rate as mediated by the protein GST-FinO141 was unexpectedly high, with a k_{app} of 1.6 x 10^6 $M^{-1}s^{-1}$. This rate represents a 19-fold increase in the rate of duplex formation as compared to the negative controls, and an ability to promote duplex formation that is 5-fold greater than the wild type protein. Recent work by Lori Jerome has demonstrated that the major RNase E cleavage site in FinP is located within the spacer region (Jerome, unpublished). I propose that the spacer is protected from RNase E degradation by the C-terminal α-helical domain (which is absent in GST-FinO141) by steric hindrance. The spacer region of antisense RNAs are known to play an important role in the formation of sense – antisense hybrids (Nordstrom and Wagner, 1994). As a result, the proximity of the negatively charged C-terminus of wild type FinO and the spacer region of FinP/TraJ may result in an inhibition of spacer/spacer interactions, and this may cause a decrease in the rate of hybridization (rate k_2). For this reason GST-FinO141 may demonstrate a greater ability to promote the overall rate of duplex formation (k_{app}) due to the absence of the C-terminal helix. Alternatively, the central region of the protein (residues 74 – 141) may play a critical positive role in duplex formation. Several possible modes of activity of the central region are proposed. First, the putative secondary RNA binding domain (residues 117 – 122) may play a role in the recognition of the incoming complementary transcript. In this model each FinO protein would bind to a FinP

or TraJ RNA *via* the N-terminal primary binding domain. The secondary binding domain would then interact with a complementary transcript that is not involved in a kissing reaction. This would cause the incoming transcript to "stall" near the loops of the primary transcript, increasing the probability of complex formation. Second, the secondary binding domain may contribute to the stability of existent kissing complexes in a Rom-like manner. In this case the primary binding domain of FinO would interact with uncomplexed transcripts in a manner reminiscent of p53, but the secondary binding domain would only interact with loops that are hybridized in a kissing complex. In this model, sequence differences between FinP alleles would not affect secondary binding by the FinO protein because binding would occur in a structurally dependent manner. Third, the central region (residues 74 – 141) might alter the RNA structure, resulting in a conformation that favors duplex formation. This would increase the rate of deep-kissing complex formation. Fourth, the central domain may be important for protein-protein interactions between FinO molecules. Dimerization between FinO molecules would bring bound RNAs into closer proximity, increasing the probability of kissing between the bound transcripts. However FinO-FinO interactions would need to be transient to allow for dissociation in the event that the bound transcripts do not form a kissing complex, i.e. both FinO proteins are bound to FinP.

It is possible that the deletion of the C-terminus could allow any of these putative mechanisms to work more efficiently due to freedom from folding constraints and/or steric hindrances. Finally, the residue C142 has been removed in the mutant protein GST-FinO141. It is possible that the deletion of the cysteine frees the central region of the protein from folding constraints. In future studies of the FinO protein site-directed mutagenesis of C142 within the context of full length GST-FinO should be done to determine the importance of this residue.

Chapter 5. Discussion

A. The Role of Downstream Sequences in Fertility Inhibition

Even though the protein responsible for fertility inhibition of F has been identified and characterized (van Biesen and Frost, 1994) it remained a possibility that the downstream sequences, including *orf86*, could play a role in fertility inhibition. To address this possibility, and to deduce the minimum *finO* sequence required for fertility inhibition, nested deletions of the 4.0 kb *Pst*I fragment from R6-5 were produced. It is evident from mating assays and the TraJ-β-galactosidase fusion reporter system that deletions of sequences downstream of *finO* do not play a role in fertility inhibition. Furthermore, even minor deletions of the *finO* coding region resulted in the derepression of mating, and the deregulation of TraJ translation, implying that the C-terminal domain of the protein is critical for FinO function *in vivo*.

The disputed "FinO as an antisense RNA" model as proposed by Walter Dempsey (McIntire and Dempsey, 1987) claims that the 21 kDa FinO protein does not play a role in R100 fertility inhibition, and that two transcripts from the opposite strand account for fertility inhibition. This model has been challenged quite vigorously by observations that the natural mutant R100-1, which contains an insertion of a single base is derepressed for transfer (Yoshioka *et al.*, 1987; Sugino and Hirota, 1962). The fact that a frameshift mutation had such a profound effect upon fertility inhibition strongly suggested that the fertility product is proteinaceous in nature. *In vitro* evidence that the 21 kDa FinO protein does directly bind FinP and furthermore protects FinP RNA from degradation by RNase E also argue against the Dempsey model (van Biesen and Frost, 1994; Jerome, unpublished). Finally, the results presented here show that, unless both proposed counter transcripts are transcribed from within the *finO* coding region, only a proteinaceous FinO product can explain the mechanism of fertility inhibition in R100 and the F-like plasmids.

B. RNA Binding Activity of FinO

Several GST-fusion mutants were produced with the aim of defining the RNA binding domain of FinO. Because several RNA binding proteins contain binding domains which are basic in nature, the pI values of various portions of the FinO protein were predicted. The possible secondary structure of the protein was also predicted. Mutants were produced on the basis of the computer-predicted characteristics of the protein.

The data from the binding studies showed that GST-FinO has an association constant of 5.6×10^7 M^{-1} for FinP RNA, which is higher than the previously published value of 4.0×10^5 M^{-1} (van Biesen and Frost, 1994). The difference in K_a values is likely due to the degree of protein damage caused by the purification methods. GST-FinO141, a deletion mutant that lacks the C-terminal predicted α-helical domain, displayed a binding affinity which was comparable to that of the full sized protein, GST-FinO. Some aggregation of the protein occurred, resulting in the formation of RNA-protein aggregates. The smaller fusion peptide, GST-FinO73 also formed RNA-protein aggregates. In the case of GST-FinO73, a smeared trail of FinP beneath the aggregates was observed, suggesting that there is a greater degree of dissociation of FinP from the GST-FinO73 protein-RNA aggregates than is seen with GST-FinO141. It is possible that the C-terminal region of the protein plays a role in stabilizing the wild type conformation of the protein. Deletion of the region, while not affecting binding *per se*, does change the migration pattern of the protein on a polyacrylamide gel due to aggregate formation. Larger deletions of FinO may have removed an alternate RNA binding motif, or may have destabilized the protein to a greater degree, resulting in the dissociation pattern observed with GST-FinO73. A similar observation was made with Tat protein. In this case the minimum, arginine-rich binding domain is well defined. When the domain was synthesized as a 9 amino acid peptide and incubated with the target RNA, TAR, a migration shift of the TAR element on polyacrylamide gels was observed (Cordingley *et al.*, 1990; Calnan *et al.*, 1991a).

However, the K_a of the 9 amino acid peptide could be increased by 2 orders of magnitude by the reintroduction of the C-terminal domain of the protein (Tao and Frankel, 1993). The authors concluded that the C-terminal domain plays an important role in stabilizing the overall structure of the protein. Further work showed that a secondary RNA binding site is located within the C-terminal domain of Tat (Loret *et al.*, 1992). A second protein, Rev, also demonstrates the importance of flanking residues upon RNA binding activity. In this case the deletion of flanking residues resulted in a change in the structure of the α-helical RNA binding domain. Chemical modifications which constrained the terminal residues of the helical region or the reintroduction of neighboring residues restored the binding affinity of the region (Daly *et al.*, 1995).

The GST fusion peptide, GST-FinO34 was not able to cause a mobility shift of FinP. This suggests that, while the first 34 residues of the protein may play an undefined role in RNA binding, they are not able to bind FinP in the absence of the N-terminal helical domain. Although important for FinO protection functions *in vivo* (see below) the C-terminal helical domain is likewise unable to bind the FinP transcript when isolated from the rest of the protein. It is concluded that the primary RNA binding site of FinO is located within the N-terminal 73 amino acids.

C. RNA Protection Activity of FinO Protein

The degree of fertility inhibition conferred by FinO deletion mutants was measured by mating assays. The degree of TraJ translation was measured indirectly by β-galactosidase assays of a *traJ-lacZ* fusion transcript when the *finO* mutants were supplied in *trans*. All *finO* mutants were found to be derepressed for conjugal transfer and deregulated for TraJ-β-galactosidase protein synthesis. It was possible that *finO* mutants were allowing for the derepression of mating/deregulation of TraJ-β-galactosidase protein expression by a number of means. First the mutant proteins may be unable to bind the transcripts. Second the proteins may be unable to protect the FinP RNA from degradation by RNases.

Finally, the proteins may be unable to promote duplex formation between the FinP and TraJ transcripts, allowing for ribosomal loading to the TraJ RBS. The plasmids encoding GST-FinO fusion protein or GST-FinO mutants that had been constructed for the purpose of defining the minimum RNA binding domain were tested for their ability to stabilize FinP *in vivo*. The results indicated that GST-FinO was able to protect FinP from degradation, while all GST-FinO derived proteins were unable to protect FinP from degradation *in vivo*. From this it is concluded that the removal of C-terminal amino acids result in the total abrogation of the FinO protection function, regardless of the binding ability of the protein. *In vitro* and *in vivo* studies by Jerome (unpublished) show that the RNase responsible for FinP degradation is RNase E, and that enzymatic cleavage occurs at two sites; one within the spacer region, the other within the 3' tail. FinO protein blocks RNase cleavage at these sites (Jerome, unpublished). From these data it is proposed that the negatively charged C-terminal α-helical domain of FinO may interact with FinP by directly binding the bases of the single-stranded spacer region, thereby protecting it from RNase E cleavage. This would differ from the mechanism by which the negatively charged helical domain of Rom operates. In this case the helical domain presses the RNA target complex into the positively charged helix via repulsive electrostatic interactions (Predki *et al.*, 1995). Although the spacer region is highly conserved among the FinP alleles (Finlay *et al.*, 1986), it is unclear at this time whether the putative protein:RNA interaction would be sequence-specific. It is unlikely that the C-terminal domain would bind to or undergo electrostatic interactions with regions of fully duplexed RNA because the negative charge of the helix would not interact favorably with the highly negatively charged phosphate backbones.

D. Promotion of Duplex Formation by FinO

GST-FinO was shown previously to increase the rate of duplex formation between FinP and TraJ by approximately 5-fold *in vitro* (van Biesen and Frost, 1994). The data presented here show a similar increase in the rate of duplex

formation, 3.7-fold, as mediated by the GST-FinO fusion. The peptide GST-FinO73 was able to increase the rate k_{app} of the antisense-sense pair at nearly GST-FinO levels, suggesting that the very act of FinO binding is able to confer an improved rate of duplex formation between FinP and TraJ. Several RNA binding proteins have been studied in very fine detail. One of these, Tat protein of HIV, binds the target RNA, TAR, and changes the bend angle of the TAR from ~25° to less than 10° (Zacharias and Hagerman, 1995). In the case of Rev/RRE binding, also from HIV, the RNA conformation is changed from the A-form to the B-form (Auer *et al.*, 1994). It seems likely, based on the binding profiles of these RNA binding proteins, that the N-terminus of FinO could mediate a change in the conformation of FinP or TraJ following binding. The mutant GST-FinO141, which lacks the C-terminal helical domain and the residue C142, endows a 19-fold increase in the rate of duplex formation between TraJ and FinP RNAs. This surprising result suggests that the duplex formation function must be provided by more than one domain of the protein. It is observed that the minimum binding domain is able to provide an increase in the rate of duplex formation. The central region of the protein (residues 74-141) which is present in the mutant GST-FinO141 provides for an additional degree of annealing between the sense and antisense RNAs. It is possible that the central region contains a secondary RNA binding site, as is true of the Tat protein (Loret *et al.*, 1992). The most likely candidate would be the highly basic stretch of residues at positions 117-122 (HKKLRR). The secondary binding domain may function by stabilizing the FinO:RNA complex, or may act to mediate interactions between the bound RNA and incoming complementary transcripts. This latter model could explain the observation by Koraimann *et al* (1996) that duplex formation between wild type TraJ and FinP loop mutants could be rescued by the FinO protein. In the case of GST-FinO141 the secondary site may be unrestrained due to the deletion of the residue C142 and/or the absence of the C-terminal helical domain. This added degree of flexibility may allow the region to recognize and bind unbound complementary transcripts more efficiently, increasing the stabilizing effect of the protein during the kissing reaction. Alternately the central region may play a role

in RNA folding, allowing for greater sequence interaction between the transcripts. Finally, the region may be important for an as yet uncharacterized protein-protein interaction, which could increase the probability of kissing, or could act to stabilize existing kissing complexes.

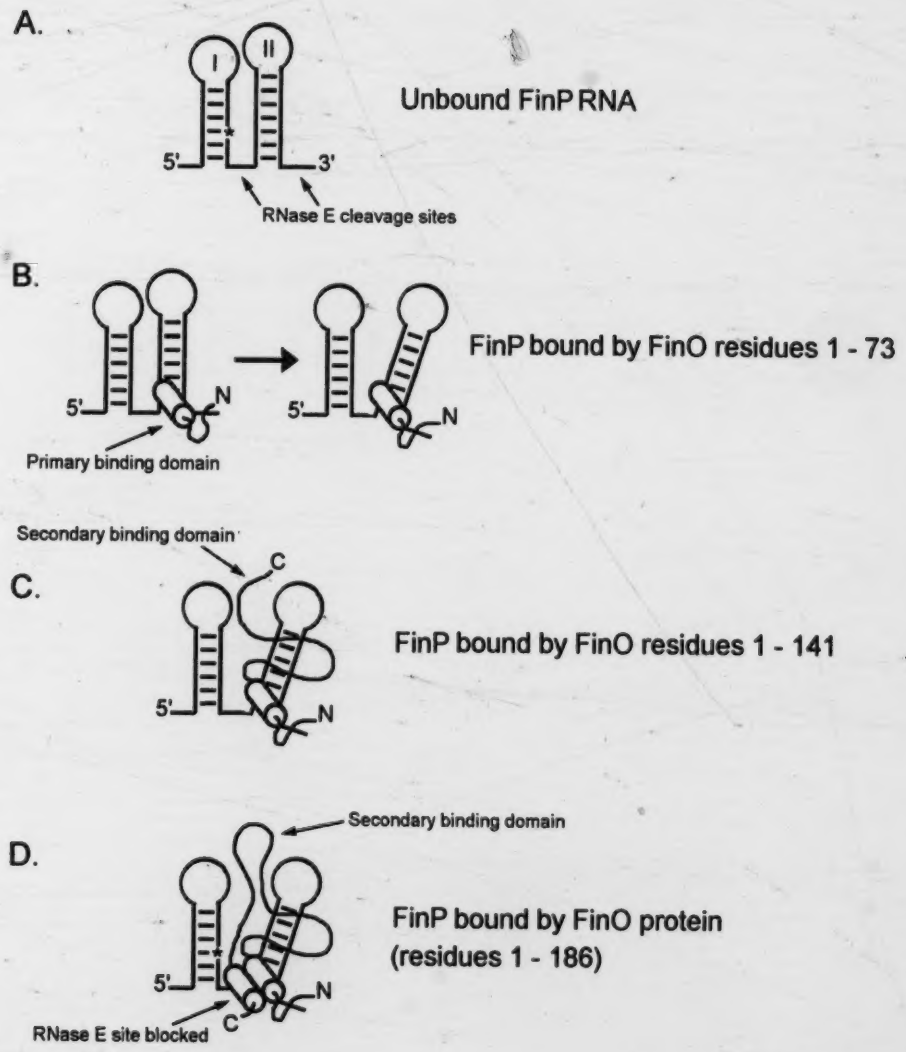
In the case of GST-FinO, the annealing potential (the ability to promote duplex formation) of the protein is approximately 5-fold lower than the mutant protein GST-FinO141. It is proposed in this communication that the C-terminal helical domain protects the FinP spacer from RNase E degradation. While critical for FinP stability *in vivo*, the protection afforded by the C-terminus may interfere with hybridization reactions between the FinP spacer and the TraJ spacer, resulting in an overall decrease in k_{app}. The spacer regions of other antisense systems are known to play an important role in the formation of their respective sense-antisense hybrids (Nordstrom and Wagner, 1994).

E. A Model of the Functional Domains of FinO

The results presented here agree well with the model of FinOP interaction suggested by van Biesen and Frost (1994). First, the presence of FinO results in the stabilization of FinP, making it available for interaction with the *traJ* mRNA. Second, FinO acts to increase the rate of duplex formation between FinP and TraJ, resulting in the formation of complexes from which TraJ protein cannot be translated (van Biesen and Frost, 1994). From the results presented here I propose the following model for the function of the various domains of the FinO protein. The model is shown diagrammatically in Figure 5.1. The N-terminal region of the protein, which contains the primary RNA binding domain, binds the base and the 3' tail of stem-loop II in a moderately specific manner. The N-terminal region of the protein covers the 3' tail, protecting it from RNase E degradation. The very act of protein binding causes a conformational change in the structure of the RNA, which results in an increase in the rate of duplex

Figure 5.1 A schematic diagram of the proposed model for the activities of the domains of the protein FinO. Panel A depicts the FinP secondary structure as previously defined. Two RNase E cleavage sites have been identified and their locations are indicated in the top figure. An additional cleavage site is located at nucleotide 30 in the FisP mutant. This site is indicated by the asterisk in the stem-loop I structure. Panel B depicts the N-terminal portion of FinO, residues 1 through 73, binding to the FinP RNA. It is believed that the FinP transcript is bent or folded by the protein, increasing the availability of the spacer region for duplex formation. The α-helical domains of the protein are portrayed as cylinders. Panel C shows binding by the truncation mutant FinO141. It is hypothesized that the central region of the protein contains a secondary binding domain which is important for binding incoming complementary transcripts. The presence of additional residues may also stabilize the N-terminal α-helical domain. Alternatively the central region may impose additional conformational changes on the transcript. Panel D shows the full sized FinO protein bound to FinP. The C-terminal helical domain binds the spacer region directly, decreasing the availability of the single-stranded region for duplex formation. The protein does not physically overlap the FisP site at U30, so does not confer protection from RNases at this site.

Figure 5.1


A.
I
II
5'
3'
RNase E cleavage sites
Unbound FinP RNA
B.
5'
N
Primary binding domain
5'
N
FinP bound by FinO residues 1 - 73
Secondary binding domain
C.
C
5'
N
FinP bound by FinO residues 1 - 141
D.
Secondary binding domain
5'
N
C
RNase E site blocked
FinP bound by FinO protein
(residues 1 - 186)

formation between the target RNA and its complementary transcript. Because the N-terminal peptide binds at the base of stem-loop II, it is likely that the conformational change involves the bending or unwinding of stem-loop II. This may allow greater access to the single-stranded spacer region.

The central region of the protein likely plays a role in increasing the rate of kissing. In this model the primary binding domain binds the target RNA at stem-loop II. The putative secondary RNA binding domain (residues 117 – 122) would then recognize and bind incoming complementary transcripts, which are not already involved in kissing complexes. This would cause the incoming transcript to "stall" near the loops of the primary transcript, increasing the probability of complex formation. It is also possible that the region promotes additional unwinding of stem-loop II or even stem-loop I, though it would be difficult to reconcile the latter suggestion with the observation that FinO does not protect the stem-loop I mutant, FisP, from RNase degradation at nucleotide U30. The reintroduction of the central region likely stabilizes the conformation of the N-terminal α-helical domain, resulting in the observed improvement in binding affinity and the decrease in the degree of dissociation of FinP and TraJ transcripts from the fusion protein. Because the N-terminal helical domain of FinO is considerably larger than the helical domain of Rev (Tan *et al.*, 1993), the effect of re-introducing neighboring residues is not as great.

Finally, the C-terminal region of the protein is responsible for conveying protection from RNase E cleavage to the FinP molecule at the 4 base, single-stranded spacer region. However protection by the C-terminal helix decreases the availability of the spacer region for duplex formation. This results in a decreased apparent rate of duplex formation by GST-FinO as compared to the mutant GST-FinO141 which does not have the C-terminal helical domain. Because of the observed instability of FisP in the presence of FinO (Finnegan and Willetts, 1971; van Biesen, 1994; Jerome, unpublished), it is suggested that the C-terminus of FinO does not interact directly with stem-loop I.

F. Direction of Further Studies

Though the region of FinO responsible for RNA binding has been identified, the minimal binding sequence has not yet been defined. Toward this goal, peptide fragments of the N-terminal end of varying sizes should be synthesized and examined for their binding proficiency. Site-directed mutagenesis of the N-terminal helical domain may uncover a wealth of information as to how the domain binds the target RNAs by identifying the critical amino residues. The "alanine scanning" technique is the best mechanism for site directed mutagenesis of the FinO helical domain due to the natural tendency of alanines to form α-helical structures (Cunningham and Wells, 1989). Chemical modification of specific bases of the FinP RNA may allow for the determination of the types of interactions between the amino acids of the binding domain and the RNA residues. Examples of modification would be the replacement of purines by N-methyl modified purines, or replacement by dI (Bartel *et al.*, 1991). Chemical modification of specific phosphates of the RNA by ethylation or substitution by methylphosphonates could be used to identify the sites where critical electrostatic interactions between the protein and the RNA occur (Churcher *et al.*, 1993; Calnan *et al.*, 1991a).

Finally, the tertiary structure of both FinO and the RNA target could be determined, either by the method of NMR or X-ray crystallography. Based on previous attempts to define the tertiary structure of the Tat RNA binding domain and the TAR RNA, it would not be surprising to find that one or both of the FinO and FinP RNA structures are relatively unstructured when unbound (Tan and Frankel, 1992; Puglisi *et al.*, 1992). Based on the same prior experiments however, it is likely that the bound complex would be highly structured. The use of such powerful tools for defining the structure of the bound complex will also provide for a much clearer understanding of the molecular basis by which FinO mediates RNA protection and increases the rate of duplex formation.

Chapter 6. References

References

Achtman, M., Kennedy, N., and Skurray, R. (1977) Cell-cell interactions in conjugating *Escherichia coli*: role of the *traT* protein in surface exclusion. *Proc. Natl. Acad. Sci. USA* **74:** 5104-5108.

Achtman, M., Willetts, N., and Clark, A. J. (1971) Beginning a genetic analysis of conjugational transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. *J. Bacteriol.* **106:** 529-538.

Achtman, M., Manning, P. A., Kusecek, B., Schwuchow, S., and Willetts, N. (1980) A genetic analysis of F sex factor cistrons needed for surface exclusion in *Escherichia coli*. *J. Mol. Biol.* **138:** 779-795.

Anderson, E. S. (1968) The ecology of transferable drug resistance in the enterobacteria. *Annual Rev. Microb.* **22:** 131-180.

Auer, M., Gremlich, H. U., Seifert, J. M., Daly, T. J., Parslow, T. G., Casari, G., and Gstach, H. (1994) Helix-loop-helix motif in HIV-1 Rev. *Biochem.* **33:** 2988-2996.

Banner, D. W., Kokkinidis, M., and Tsernoglou, D. (1987) Structure of the ColE1 Rop protein at 1.7 Angstrom resolution. *J. Mol. Biol.* **196:** 657-675.

Bartel, D. P., Zapp, M. L., Green, M. R., and Szostak, J. W. (1991) HIV-1 rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **67:** 529

Battiste, J. L., Mao, H., Rao, N. S., Tan, R., Muhandiram, D. R., Kay, L. E., Frankel, A. D., and Williamson, J. R. (1996) α-helix-RNA major groove

recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **273:** 1547-1551.

Berkhout, B., and Jeang, K. T. (1989) *Trans* activation of human immunodeficiency virus type 1 is sequence specific for both the single-stranded bulge and loop of the *trans*-acting-responsive hairpin: a quantitative analysis. *J. Virol.* **63:** 5501.

Blomberg, P., Wagner, E. G. H., and Nordstrom, K. (1990) Control of replication of plasmid R1: the duplex between the antisense RNA, CopA, and its target, CopT, is processed specifically *in vivo* and *in vitro* by RNase III. *EMBO J.* **9:** 2331-2340.

Brantl, S., and Behnke, D. (1992) Copy number control of the streptococcal plasmid pIP501 occurs at three levels. *Nucl. Acids Res.* **20:** 395-400.

Brantl, S., and Wagner, E. G. H. (1994) Antisense RNA-mediated transcriptional attenuation occurs faster than stable antisense/target RNA pairing: an *in vitro* study of plasmid pIP501. *EMBO J.* **13:** 3599-3607.

Brantl, S., Birch-Hirschfeld, E., Behnke, D. (1993) RepR protein expression on plasmid pIP501 is controlled by an antisense RNA-mediated transcription attenuation mechanism. *J. Bacteriol.* **175:** 4052-4061.

Calnan, B. J., Biancalana, S., Hudson, D., and Frankel, A. D. (1991a) Analysis of arginine-rich peptides from the HIV Tat protein reveals unusual features of RNA-protein recognition. *Genes Dev.* **5:** 201-210.

Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D., and Frankel, A. D. (1991b) Arginine-mediated RNA recognition: the arginine fork. *Science* **252:** 1167.

Carter, J. R., and Porter, R. D. (1991) *traY* and *traI* are required for *oriT*-dependent enhanced recombination between *lac*-containing plasmids and *λplac5*. *J. Bacteriol.* **173:** 1027-1034.

Case, C. C., Simons, E. L., and Simons, R. W. (1990) The IS*10* transposase mRNA is destabilized during antisense RNA control. *EMBO J.* **9:** 1259-1266.

Case, C. C., Roels, S. M., Jensen, P. D., Lee, J., Kleckner, N., and Simons, R. W. (1989) The unusual stability of the IS*10* anti-sense RNA is critical for its function and is determined by the structure of its stem-domain. *EMBO J.* **8:** 4297-4305.

Casabadan, M.J., and Cohen, S. N. (1980) *J. Mol. Biol.* **138:** 179-207.

Chandler, M., and Galas, D. (1983) Cointegrate formation mediated by Tn9 II. Activity of IS1 is modulated by external DNA sequences. *J. Mol. Biol.* **170:** 61-91.

Chang, A. C. Y., and Cohen, S. N. (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the p15a cryptic miniplasmid. *J. Bacteriol.* **143:** 1141-1156.

Chang, A. -Y., and Tinoco, I. (1997) The structure of an RNA "kissing" hairpin complex of the HIV TAR hairpin loop and its complement. *J. Mol. Biol.* **269:** 52-66.

Chapman, J. S., and Carlton, B. C. (1985) Conjugal plasmid transfer in *Bacillus thuringiensis*. *In* Helinski, D. R., Cohen, S. N., Clewell, D. B., Jackson, D. A., and Hollaender, A. (eds). *Plasmids in Bacteria*. pp. 453-467. Basic Life Sciences, Plenum, New York.

Cheah, K. C., and Skurray, R. A. (1986) The F plasmid carries an IS*3* insertion within *finO*. *J. Gen. Microbiol.* **132:** 3269-3275.

Churcher, M. J., Lamont, C., Hamy, F., Dingwall, C., Green, S. M., Lowe, A. D., Butler, J. G., Gait, M. J., and Karn, J. (1993) High affinity binding of TAR RNA by the human immunodeficiency virus type-1 Tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region. *J. Mol. Biol.* **230:** 90-110.

Cook, K. S., Fisk, G. J., Hauber, G. J., Usman, N., Daly, T. J., and Rusche, J. R. (1991) Characterization of HIV-1 Rev protein: binding and stoichiometry and minimal RNA substrate. *Nucleic Acid Res.* **19:** 1577.

Cordingley, M. G., LaFemina, R. L., Callahan, P. L., Condra, J. H., Sardana, V. V., Graham, D. J., Nguyen, T. M., LeGrow, K., Gotlib, L., Schlabach, A. J., and Colonno, R. J. (1990) Sequence specific interaction of tat and tat peptides with the TAR sequence of HIV-1 *in vitro*. *Proc. Natl. Acad. Sci. USA.* **87:** 8985.

Cosma, C. L., Danese, P. N., Carlson, J. H., Silhavy, T. J., and Snyder, W. B. (1995) Mutational activation of the Cpx signal transduction pathway of Escherichia coli suppresses the toxicity conferred by certain envelope-associated stresses. *Mol. Microbiol.* **18:** 491-505.

Cunningham, B. C., and Wells, J. A. (1989) Higher resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **244:** 1081-1085.

Cuozzo, M., and Silverman, P. M. (1986) Characterization of the F plasmid TraJ protein synthesized in F' and Hfr strains of *Escherichia coli* K-12. *J. Biol. Chem.* **261:** 5175-5179.

Daly, T. J., Doten, R. C., Rusche, J. R., and Auer, M. (1995) The amino terminal domain of HIV-1 Rev is required for discrimination of the RRE from nonspecific RNA. *J. Mol. Biol.* **253:** 243-258.

Daly, T. J., Cook, K. S., Gary, G. S., Maione, T. E., and Rusche, J. R. (1989) Specific binding of HIV-1 recombinant rev protein to the *rev*-response element *in vitro*. *Nature* **342:** 816.

Daly, T. J., Doten, R. C., Rennert, P., Auer, M., Jaksche, H., Donner, A., Fisk, G., and Rusche, J. R. (1993a) Biochemical characterization of binding of multiple HIV-1 Rev monomeric proteins to the Rev responsive element. *Biochem.* **32:** 10497-10505.

Daly, T. J., Rennert, P., Lynch, P., Barry, J. K., Dundas, M., Rusche, J. R., Doten, R. C., Auer, M., and Farrington, G. K. (1993b) Perturbation of the carboxy terminus of HIV-1 Rev affects multimerization on the Rev responsive element. *Biochem.* **32:** 8945-8954.

Datta,N. (1975) Epidemiology and classification of plasmids. *In* Schlessinger, D. (ed.) *Microbiology – 1974.* pp. 9-15. American Society for Microbiology, Washington, D.C.

Delling, U., Roy, S., Sumner-Smith, M., Barnett, R., Reid, L., Rosen, C. A., and Sonenberg, N. (1991) The number of positively charged amino acids in the basic domain of Tat is critical for *trans*-activation and complex formation with TAR RNA. *Proc. Natl. Acad. Sci. USA.* **88:** 6234-6238.

Dempsey, W. B. (1987) Transcript analysis of the plasmid R100 *traJ* and *finP* genes. *Mol. Gen. Genet.* **209:** 533-544.

Dempsey, W. B. (1989) Sense and antisense transcripts of *tra*M, a conjugal transfer gene of the antibiotic resistance plasmid R100. *Mol. Micro.* **3:** 561-570.

Dempsey, W. B. (1993) Key regulatory aspects of transfer of F related plasmids. *In* Clewell, D. B. (ed.) *Bacterial Conjugation.* pp. 53-73. Plenum Press, New York

Dempsey, W. B. (1994) *traJ* sense RNA initiates at two different promoters in R100-1 and forms two stable hybrids with antisense *fin*P RNA. *Mol. Micro.* **13:** 313-326.

Dempsey, W. B., and Fee, B. E. (1990) Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. *Mol. Micro.* **4:** 1019-1028.

Di Laurenzio, L., Finlay, B. B., Frost, L. S., and Paranchych, W. (1991) Characterization of the *oriT* region of the IncFV plasmid pED208. *Mol. Microbiol.* **5:** 1779-1790.

Eberle, W., Klaus, W., Cesareni, G., Sander, C., and Rosch, P. (1990) Proton nuclear magnetic resonance assignments and secondary structure of the ColE1 rop (rom) protein. *Biochem.* **29:** 7402-7407.

Eguchi, Y., and Tomizawa, J. (1990) Complex formed by complementary RNA stem-loops and its stablilization by a protein: function of ColE1 Rom protein. *Cell* **60:** 199-209.

Eguchi, Y., Itoh, T., and Tomizawa, J. (1991) Antisense RNA. *Annu. Rev. Biochem.* **60:** 631-652.

Farrand, S. K. (1993) Conjugal transfer of *Agrobacterium* plasmids. *In* Clewell, D. B. (ed). *Bacterial Conjugation*. pp. 255-285. Plenum Press, Plenum, New York.

Fee, B. E., and Dempsey, W. B. (1986) Cloning, mapping and sequencing of plasmid R100 *tra*M and *fin*P genes. *J. Bacteriol.* **167:** 336-345.

Finlay, B. B., Frost, L. S., Paranchych, W., and Willetts, N. S. (1986) Nucleotide sequences of five IncF plasmid *fin*P alleles. *J. Bacteriol.* **167:** 754-757.

Finnegan, D. J., and Willetts, N. S. (1971) Two classes of F *lac* mutants insensitive to transfer inhibition by an F-like R factor. *Mol. Gen. Genet.* **111:** 256-264.

Firth, N., and Skurray, R. (1992) Characterization of the F plasmid bifunctional conjugation gene, *traG*. *Mol. Gen. Genet.* **232:** 145-153.

Firth, N., Ippen-Ihler, K., and Skurray, R. A. (1996) Structure and function of the F factor and mechanism of conjugation. *In* Neidhardt, F. C. (ed.) *Escherichia coli and Salmonella,* (2nd Edition). American Society for Microbiology, Washington, D.C.

Fowler, T., and Thompson, R. (1986) Shadow promoters in the F plasmid transfer operon. *Mol. Gen. Genet.* **202:** 509-511.

Frangioni, J.V., and Neel, B.G. (1993) Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal. Biochem.* **210:** 179-187.

Frankel, A. D. (1995) Using peptides to study RNA-protein recognition. *In.* Nagai, K., and Mattaj, I. W. (eds.) *RNA-Protein Interactions.* pp.221-247. Oxford University Press, New York.

Friedman, P.N., Chen, X., Bargonetti, J., and Prives, C. (1993) The p53 protein is an unusually shaped tetramer that binds directly to DNA. *Proc Natl Acad Sci USA.* **90:** 3319-3323.

Freifelder, D. (1968) Studies with *E. coli* sex factors. *Cold Spring Harbor Symp. Quant. Biol.* **33:** 425-434.

Frost, L. S., Ippen-Ihler, K., and Skurray, R. A. (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Micro. Revs.* **58:** 162-210.

Gaudin, H. M., and Silverman, P. M. (1993) Contributions of promoter context and structure to regulated expression of the F plasmid *traY* promoter in *Escherichia coli* K-12. *Mol. Microbiol.* **8:** 335-342.

Graus-Goldner, A., Graus, H., Schlacher, T., and Hogenauer, G. (1990) The sequences of genes bordering *oriT* in the enterotoxin plasmid P307: comparison with the sequences of plasmids F and R1. *Plasmid* **24:** 119-131.

Gregorian, R. S., and Crothers, D. M. (1995) Determinants of RNA hairpin loop-loop complex stability. *J. Mol. Biol.* **248:** 968-984.

Hamy, F., Asseline, U., Grasby, J., Iwai, S., Pritchard, C., Slim, G., Bulter, P. J. G., Karn, J., and Gait, M. J. (1993) Hydrogen-bonding contacts in the major groove are required for human immunodeficiency virus type-1 tat protein recognition of TAR RNA. *J. Mol. Biol.* **230:** 111.

Hanahan, D. (1983) Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol.* **166:** 557-580.

Harrington, L. C., and Rogerson, A. C. (1990) The F pilus of Escherichia coli appears to support stable DNA transfer in the absence of wall-to-wall contact between cells. *J. Bacteriol.* **172:** 7263-7264.

He, L., Soderbom, F., Wagner, E. G. H., Binnie, U., Binns, N., and Masters, M. (1993) PcnB is required for the rapid degradation of RNA I, the antisense RNA that controls the copy number of ColE1-related plasmids. *Mol. Microbiol.* **9:** 1131-1142.

Heaphy, S., Finch, J. T., Gait, M. J., Karn, J., and Singh, M. (1991) Human immunodeficiency virus type-1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich 'bubble' located within the rev-responsive region of viral RNA. *Proc. Natl. Acad. Sci. USA.* **88:** 7366.

Heaphy, S., Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Karn, J., Lowe, A. D., Singh, M., and Skinner, M. A. (1990) HIV-1 regulator of virion expression (rev) protein binds to an RNA stem-loop structure located within the *rev*-response element region. *Cell* **60:** 685.

Huang, X., Hope, T. J., Bond, B. L., McDonald, D., Grahl, K., and Parslow, T. G. (1991) Minimal Rev-response element for Type 1 human immunodeficiency virus. *J. Virol.* **65:** 2131.

Inamoto, S., Yoshioka, Y., and Ohtsubo, E. (1988) Identification and characterization of the products from the *traJ* and *traY* genes of plasmid R100. *J. Bacteriol.* **170:** 2749-2757.

Ippen-Ihler, K., and Maneewannakul, S. (1991) Conjugation among enteric bacteria: mating systems dependent on expression of pili. *In* Dworkin, M. (ed.) *Microbial Cell-Cell Interactions*. pp. 35-69. American Society for Microbiology, Washington D.C.

Ippen-Ihler, K., and Skurray, R. A. (1993) Genetic organization of transfer-related determinants on the sex factor F and related plasmids. *In* Clewell, D. (ed.) *Bacterial Conjugation*. pp23-52. Plenum Press, New York.

Jain, C. (1995) IS*10* antisense control *in vivo* is affected by mutations throughout the region of complementarity between the interacting RNAs. *J. Mol. Biol.* **246:** 5585-5594.

Jakobovits, A., Smith, D. H., Jakobovits, E. B., and Capon, D. J. (1988) A discrete element 3' of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activiation by an HIV *trans* activator. *Mol Cell Biol.* **8:** 2555.

Jerome, L., unpublished.

Johnson, D., and Willetts, N. S. (1980) Construction and characterization of multicopy plasmids containing the entire F transfer region. *Plasmid* **4:** 292-304.

Kern, S. E., Kinzler, K. W., Bruskin, A., Jarosz, D., Friedman, P., Prives, C., and Vogelstein, B. (1991) Identification of p53 as a sequence-specific DNA-binding protein. *Science* **252:** 1708-1711.

Kittle, J. D., Simons, R. W., Lee, J., and Kleckner, N. (1989) Insertion sequence IS*10* anti-sense pairing initiates by an interaction between the 5' end of the target RNA and a loop in the anti-sense RNA. *J. Mol. Biol.* **210:** 561-572.

Kjems, J., and Sharp, P. A. (1993) The basic domain of Rev from human immunodeficiency virus type 1 specifically blocks the entry of U4/U6.U5 small nuclear ribonucleoprotein in spliceosome assembly. *J. Virol.* **67:** 4769-4776.

Kjems, J., Calnan, B. J., Frankel, A. D., and Sharp, P. A. (1992) Specific binding of a basic peptide from HIV-1 Rev. *EMBO J.* **11:** 1119.

Kleckner, N. (1989) Transposon Tn10. *In* Berg, D., and Howe, M (eds). *Mobile DNA,* pp 227-268. American Society for Microbiology, Washington, D. C.

Klimke, W. A., Fekete, R., Manchak, J., Anthony, K., and Frost, L. S. (1997) accepted to *Proceedings of the Fallen Leaf Lake Conference: Unifying Themes Among Protein and Nucleic Acid Transporters of Bacteria.*

Koraimann, G., Koraimann, C., Koronakis, V., Schlager, S., and Hogenauer, G. (1991). Repression and derepression of conjugation of plasmid R1 by wild-type and mutated *finP* antisense RNA. *Mol. Microbiol.* **5:** 77-87.

Koraimann, G., Teferle, K., Markolin, G., Woger, W., and Hogenauer, G. (1996) The FinOP repressor system of plasmid R1: analysis of the antisense RNA control of *traJ* expression and conjugative DNA transfer. *Mol. Microbiol.* **21:** 811-821.

Koronakis, V. E., Bauer, E., and Hogenauer, G. (1985) The traM gene of the resistance plasmid R1: comparison with the corresponding sequence of the *Escherichia coli* F factor. *Gene* **36:** 79-86.

Lane, H. E. D. (1981) Replication and incompatibility of F and plasmids in the IncFI group. *Plasmids* **5:** 100-126.

Laspia, M. F., Rice, A. P., and Mathews, M. B. (1989) HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. *Cell* **59:** 283-292.

Le, S. Y., Pattabiraman, N., and Maizel, J. V. Jr. (1994) RNA tertiary structure of the HIV RRE domain II containing non-Watson-Crick base pairs GG and GA: molecular modeling studies. *Nucl. Acids Res.* **22:** 3966-3976.

Lederberg, J., and Tatum, E. (1946) Gene recombination in *E. coli*. *Nature* **158:** 558.

Lee, S. H., Frost, L. S., and Paranchych, W. (1992) FinOP repression of the F plasmid involves extension of the half-life of FinP antisense RNA by FinO. *Mol. Gen. Genet.* **235:** 131-139.

Levin, B. R., and Lenski, R. E. (1983) Coevolution in bacteria and the viruses and plasmids. *In.* Futuyma and Slatkin (eds). *Coevolution.* pp. 99-127. Sinauer Assoc., Sunderland, Massachusetts.

Liu, Y., Wang, Z., and Rana, T. M. (1996) Visualizing a specific contact in the HIV-1 Tat protein fragment and trans-activation response region RNA complex by photocross-linking. *J. Biol. Chem.* **271:** 10391-10396.

Loret, E. P., Georgel, P., Johnson, W. C. Jr., Ho, P. S. (1992) Circular dichroism and molecular modeling yield a structure for the complex of human immunodeficiency virus type 1 *trans*-activation response RNA and the binding region of Tat, the *trans*-acting transcriptional activator. *Proc. Natl. Acad. Sci. USA.* **89:** 9734-9738.

Lundquist, P. D., and Levin, B. R. (1986) Transitory derepression and the maintenance of conjugative plasmids. *Genetics* **113:** 483-497.

Ma, C., and Simons, R. W. (1990) The IS*10* antisense RNA blocks ribosome binding at the transposase translation initiation site. *EMBO J.* **9:** 1267-1274.

Madore, S. J., Tiley, L. S., Malim, M H., and Cullen, B. R. (1994) Sequence requirements for Rev multimerization *in vivo*. *Virology* **202:** 186-194.

Malim, M. H., Bohnlein, S., Hauber, J., and Cullen, B. R. (1989) Functional dissection of the HIV-1 Rev trans-activator--derivation of a *trans*-dominant repressor of Rev function. *Cell 58:* 205-214

Maneewannakul, K., Maneewannakul, S., and Ippen-Ihler, K. (1992) Sequence alterations affecting F plasmid transfer gene expression: a conjugation system dependent on transcription by the RNA polymerase of phage T7. *Mol. Microbiol.* **6:** 2961-2973.

Marmur, J., Rownd, R., Falkow, W., Baron, L.S., Schildkraut, C., and Doty, P. (1961) The nature of intergeneric episomal infection. *Proc. Natl. Acad. Sci. USA* **47:** 972-979.

Masai, H., and Arai, K. I. (1988) *repA* protein and *oriR*-dependent initiation of R1 plasmid replication: identification of a rho-dependent transcription terminator required for cis-action of *repA* protein. *Nucleic Acids Res.* **16:** 6493-6514.

Masai, H., Kaziro, Y., and Arai, K. I. (1983) Definition of *oriR*, the minimum DNA segment essential for initiation of R1 plasmid replication *in vitro*. *Proc. Natl. Acad. Sci. USA.* **80:** 6814-6818.

Masukata, H., and Tomizawa, J. (1986) Control of primer formation for ColE1 plasmid replication: conformational change of the primer transcript. *Cell* **44:** 125-136.

McCormick, F., Clark, R., Harlow, E., and Tjian, R. (1981) SV40 antigen binds specifically to a cellular 53 K protein *in vitro*. *Nature* **292:** 63-65.

McIntire, S. A., and Dempsey, W. B. (1987). Fertility inhibition gene of plasmid R100. *Nucl. Acids Res.* **15:** 2029-2042.

Meynell, E., and Datta, N. (1967) Mutant drug resistant factors of high transmissibility. *Nature* **214:** 885-887.

Miki, T., Horiuchi, T., and Willetts, N. S. (1978) Identification and characterization of four new *tra* cistrons on the *E. coli*. K12 sex factor F. *Plasmid* **1:** 316-323.

Mileykovskaya, E., and Dowhan, W. (1997) The Cpx two-component signal transduction pathway is activated in *Escherichia coli* mutant strains lacking phosphatidylethanolamine. *J. Bacteriol.* **179:** 1029-1034.

Miller, J. (1972) pp. 201-205. In: *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Mullineaux, P., and Willetts, N. (1984) Promoters in the transfer region of plasmid F. *In* Helinski, D., Cohen, S., Clewell, D., and Hollaender, A. (eds). *Plasmids in Bacteria.* pp. 605-614. Plenum Press, New York.

Mullineaux, P., and Willetts, N. (1985) Promoters in the transfer region of plasmid F. *Basic Life Sci.* **30:** 605-614.

Nakaya, R., Nakamura, A., and Murata, Y. (1960) Resistance transfer agents in *Shigella*. *Biochem. Res. Commun.* **3:** 654-659.

Nedbal,W., Frey, M., Willemann, B., Zentgraf, H., and Sczakiel, G. (1997) Mechanistic insights into p53-promoted RNA-RNA annealing. *J. Mol. Biol.* **266:** 677-687.

Nelson, W. C., Morton, B. S., Lahue, E. E., and Matson, S. W. (1993) Characterization of the *Escherichia coli* F factor *traY* gene product and its binding sites. *J. Bacteriol.* **175:** 221-2228.

Nordstrom, K., and Wagner, E. G. H. (1994) Kinetic aspects of control of plasmid replication by antisense RNA. *Trends Biochem Sci.* **19:** 294-300

Novick, R. P., and Hoppensteadt, F. C. (1978) On plasmid incompatibility. *Plasmid* **1:** 421-443.

Novick, R. P., Iordanescu, S., Projan, S. J., Kornblum, J., and Edelman, I. (1989) pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. *Cell* **59:** 395-404.

Oberosler, P., Hloch, P., Ramsperger, U., and Stahl, H. (1993) p53-catalyzed annealing of complementary single-stranded nucleic acids. *EMBO J.* **12:** 2389-2396.

Pavletich, N. P., and Pabo, C. O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Angstroms. *Science* **252:** 809-817.

Penfold, S. S. (1995) Thesis, University of Alberta, Edmonton.

Penfold, S. S., Simon, J., and Frost, L. S. (1996) Regulation of the expression of the *tra*M gene of the F sex factor of *Escherichia coli*. *Mol. Micro.* **20:** 549-558.

Persson, C., Wagner, E. G. H., and Nordstrom, K. (1988) Control of replication of plasmid R1: kinetics of *in vitro* interaction between the antisense RNA, CopA, and its target, CopT. *EMBO J* **7:** 3270-3288.

Persson, C., Wagner, E. G. H., and Nordstrom, K. (1990a) Control of replication of plasmid R1: structures and sequences of the antisense RNA, CopA, required for its binding to the target RNA, CopT. *EMBO J.* **9:** 3767-3775.

Persson, C., Wagner, E. G. H., and Nordstrom, K. (1990b) Control of replication of plasmid R1: formation of an initial transient complex is rate-limiting for anitsense RNA-target RNA pairing. *EMBO J.* **9:** 3777-3785.

Pogiano, J., Lynch, A. S., Bein, D., Lin, E. C. C., and Beckwith, J. (1997) Regulation of *Escherichia coli cell* envelope proteins involved in protein folding and degradation by the Cpx two-component system. *Genes Dev.* **11:** 1169-1182.

Predki, P. F., Nayak, L. M., Gottlieb, M. B. C., Regan, L. (1995) Dissecting RNA-protein interactions: RNA-RNA recognition by Rop. *Cell* **80:** 41-50.

Puglisi, J. D., Chen, L., Frankel, A. D., and Williamson, J. R. (1993) Role of RNA structure in arginine recognition of TAR RNA. *Proc. Natl. Acad. Sci. USA.* **90:** 3680-3684.

Ray, A., and Skurray, R. (1983) Cloning and polypeptide analysis of the leading region in F plasmid DNA transfer. *Plasmid* **9:** 262-272.

Riordan, F. A., Bhattacharyya, A., McAsteer, S., and Lilley, D. M. (1992) Kinking of RNA helices by bulged bases, and the structure of the human immunodeficiency virus transactivator response element. *J. Mol. Biol.* **226:** 305-310.

Rosen, C. A., Sodroski, J. G., and Haseltine, W. A. (1985) The location of *cis*-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. *Cell* **41:** 813.

Roy, S., Delling, U., Chen, C. H., Rosen, C. A., and Sonenberg, N. (1990) A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated *trans*-activation. *Genes Dev.* **4:** 1365-1373.

Saadi, S., Maas, W. K., Hill, D. F., and Bergquist, P. L. (1987) Nucleotide sequence analysis of RepFIC, a basic replicon present in IncFI plasmids P307 and F, and its relation to the RepA replicon of IncFII plasmids. *J. Bacteriol.* **169:** 1836-1846.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual, second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Scott, J. R. (1984) Regulation of plasmid replication. *Microbiol. Rev.* **48:** 1-23.

Schwab, M., Gruber, H., and Hogenauer, G. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Micro.* **5:** 439-446.

Selby, M. J., and Peterlin, B. M. (1990) Trans-activation by HIV-1 Tat via a heterologous RNA binding protein. *Cell* **62:** 769-776.

Sharp, P. A., Cohen, S. N., and Davidson, N. (1973) Electron microscope heteroduplex studies of sequence relations among plasmids of *Escherichia coli*. II. Structure of drug resistance (R) factors and F factors. *J. Mol. Biol.* **75:** 235-255.

Silhavy, T. J., Berman, M. L., and Enquist, L. W. (1984) In: *Experiments with Gene Fusions*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Silverman, P. M., Wichersham, E., and Harris, R. (1991) Regulation of the F plasmid *traY* promoter in *Escherichia coli* by host and plasmid factors. *J. Mol. Biol.* **218:** 119-128.

Silverman, P. M., Tran, L., Harris, R., and Gaudin, H. M. (1993) Accumulation of the F plasmid TraJ protein in *cpx* mutants of *Escherichia coli. J. Bacteriol.* **175:** 921-925.

Simons, R. W., and Kleckner, N. (1983) Translational control of IS*10* transposition. *Cell* **34:** 683-691.

Simonsen, L. (1990) Dynamics of plasmid transfer on surfaces. *J. Gen. Microbiol.* **136:** 1001-1007.

Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F., and Haseltine, W. A. (1985) Location of the trans-acting region on the genome of human T-cell lymphotropic virus type III. *Science* **229:**74.

Sodroski, J., Goh, W. C., Rosen, C. C., Dayton, A., Terwilliger, E., and Haseltine, W. A. (1986) A second post-transcriptional activator gene required for HTLV-III replication. *Nature* **321:** 412.

Steege, D. A. (1983) A nucleotide change in the anticodon of an Escherichia coli serine transfer RNA results in supD-amber suppression. *Nucl. Acids Res.* **11:** 3823-3832.

Steinmeyer, I. K., and Deppert, W. (1988) DNA binding properties of murine p53. *Oncogene* **3:** 501-507.

Stratagene (1995) Catalog, Stratagene Cloning Systems, USA.

Sugino, Y., and Hirota, Y. (1962) Conjugal fertility associated with resistance factor R in *Escherichia coli. J. Bacteriol.* **84:** 902-910.

Sukupolvi, S., and O'Connor, C. D. (1990) TraT lipoprotein, a plasmid-specified mediator of interactions between gram negative bacteria and their environment. *Micro. Rev.* **54:** 331-341.

Tabor, S., and Richardson, C. C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA.* **82:** 1074-1078.

Talanian, R. V., McKnight, C. J., and Kim, P. S. (1990) Sequence-specific DNA binding by a short peptide dimer. *Science* **249:** 769.

Tan, R., and Frankel, A. D. (1992) Circular dichroism studies suggest that TAR RNA changes conformation upon specific binding of arginine or guanidine. *Biochem.* **31:** 10288-10294.

Tan, R., and Frankel, A. D. (1994) Costabilization of peptide and RNA structure in an HIV Rev peptide-RRE complex. *Biochem.* **33:** 14579-14585.

Tan, R., Chen, L., Buettner, J. A., Hudson, D., and Frankel, A. D. (1993) RNA recognition by an isolated alpha helix. *Cell* **73:** 1031-1040.

Tao, J., and Frankel, A. D. (1993) Electrostatic interactions modulate the RNA-binding and transactivation specificities of the human immunodeficiency virus

and simian immunodeficiency virus Tat proteins. *Proc. Natl. Acad. Sci. USA.* **90:** 1571-1575.

Thompson, R., and Taylor, L. (1982) Promoter mapping and DNA sequencing of the F plasmid transfer genes *traM* and *traJ*. *Mol. Gen. Genet.* **188:** 513-518.

Tomizawa, J. (1984) Control of ColE1 plasmid replication: the process of binding of RNA I to the primer transcript. *Cell* **38:** 861-870.

Tomizawa, J. (1985) Control of ColE1 plasmid replication: initial interaction of RNA I and the primer is reversible. *Cell* **40:** 527-535.

Tomizawa, J. (1986) Control of ColE1 plasmid replication: binding of RNA I to RNA II and inhibition of primer formation. *Cell* **47:** 89-97.

Tomizawa, J. (1990) Control of ColE1 plasmid replication: interaction of Rom protein with an unstable complex formed by RNA I and RNA II. *J. Mol. Biol.* **212:** 695-708.

Tomizawa, J., and Itoh, T. (1981) Plasmid incompatibility determined by interaction of RNA I with primer transcript. *Proc. Natl. Acad. Sci. USA.* **78:** 6096-6100.

Tomizawa, J., and Som, T. (1984) Control of plasmid replication: enhancement of binding of RNA I to the primer transcript by the Rom protein. *Cell* **38:** 871-878.

Tomizawa, J., Itoh, T., Selzer, G., and Som, T. (1981) Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc. Natl. Acad. Sci. USA.* **78:** 1421-1425.

Tsai, M.M., Fu, Y.H. F., and Deonier, R.C. (1990) Intrinsic bends and integration host factor binding at F plasmid *ori*T. *J. Bacteriol.* **172:** 4603-4609.

United States Biochemical (1994) Life Science Research Products, USB, United States of America.

van Biesen, T. (1994) Thesis, University of Alberta, Edmonton.

van Biesen, T., and Frost, L. S. (1992) Differential levels of fertility inhibition among F-like plasmids are related to the cellular concentration of *fin*O mRNA. *Mol. Micro.* **6:** 771-780.

van Biesen, T., and Frost, L. S. (1994) The FinO protein of IncF plasmids binds FinP antisense RNA and its target, *tra*J mRNA, and promotes duplex formation. *Mol. Micro.* **14:** 427-436.

van Biesen, T., Soderbom, F., Wagner, E. G. H., and Frost, L.S. (1993) Structural and functional analyses of the FinP antisense RNA regulatory system of the F conjugative plasmid. *Mol. Micro.* **10:** 35-43.

Wagner, E. G. H., and Simons, R., W. (1994) Antisense RNA control in bacteria, phages and plasmids. *Annu. Rev. Microbiol.* **48:** 713-742.

Wagner, E. G. H., Blomberg, P., Nordstrom, K. (1992) Replication control in plasmid R1: duplex formation between the antisense, CopA, and its target, CopT, is not required for inhibition of RepA synthesis. *EMBO J.* **11:** 1195-1203.

Wagner, E. G. H., von Heijne, J., Nordstrom, K. (1987) Control of replication of plasmid R1: Translation of the 7k reading frame in the RepA mRNA leader region counteracts the interaction between CopA RNA and CopT RNA. *EMBO J.* **6:** 515-522.

Wang, Z., and Rana, T. M. (1996) RNA conformation in the Tat-TAR complex determined by site-specific photo-cross-linking. *Biochem.* **35:** 6491-6499.

Watanabe, T., and Fukasawa, T. (1961) Episome-mediated transfer of drug resistance in *Enterobacteriaceae*. I. Transfer of resistance factors by conjugation. *J. Bacteriol.* **81:** 669-678.

Weeks, K. M., and Crothers, D. M. (1991) RNA recognition by Tat-derived peptides: interaction in the major groove? *Cell* **66:** 577-588.

Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A., and Crothers, D. M. (1990) Fragments of the HIV-1 Tat protein specifically bind TAR RNA. *Science* **249:** 1281-1285.

Werstuck, G., Zapp, M. L., and Green, M. R. (1996) A non-canonical base pair within the human immunodeficiency virus rev-responsive element is involved in both rev and small molecule recognition. *Chem Biol.* **3:** 129-137.

Williamson, J. R., Battiste, J. L., Mao, H., and Frankel, A. D. (1995) Interaction of HIV Rev peptides with the Rev response element RNA. *Nucl. Acids Symp. Ser.* **114:** 46-48

Wilkins, B., and Lanka, E. (1993) DNA processing and replication during plasmid transfer between gram negative bacteria. *In* Clewell, D. B. (ed.) *Bacterial Conjugation.* pp. 105-136. Plenum Press, New York.

Willetts, N. (1977) The transcriptional control of fertility in F-like plasmids. *J. Mol. Biol.* **112:** 141-148.

Willetts, N., and Maule, J. (1986) Specificities of IncF plasmid conjugation genes. *Genet. Res.* **47:** 1-11.

Willetts, N. S., and Wilkins, B. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* 48: 24-41.

Wu, R., Wang, X., Womble, D. D., and Rownd, R. H. (1992) Expression of the *repA1* gene of IncFII plasmid NR1 is translationally coupled to expression of an overlapping leader peptide. *J. Bacteriol.* **174:** 7620-7628.

Yoshioka, Y., Fujita, Y., and Ohtsubo, E. (1990) Nucleotide sequence of the promoter-distal region of the *tra* operon of plasmid R100, including *traI* (DNA helicase I) and *traD* genes. *J. Mol. Biol.* **214:** 39-53.

Yoshioka, Y., Ohtsubo, H., and Ohtsubo, E. (1987) Repressor gene *finO* in plasmids R100 and F: constitutive transfer of plasmid F is caused by insertion of IS3 into F *finO*. *J. Bacteriol.* **169:** 619-623.

Zacharias, M., and Hagerman, P. J. (1995) The bend in RNA created by the *trans*-activation response element bulge of human immunodeficiency virus is straightened by arginine and by Tat-derived peptides. *Proc. Natl. Acad. Sci. USA.* **92:** 6052-6056.

Zapp, M. L., and Green, M. R. (1989) Sequence-specific RNA binding by the HIV-1 Rev protien. *Nature* **342:** 714-716.

Zapp, M. L., Hope, T. J., Parslow, T. G., and Green, M. R. (1991) Oligomerization and RNA binding domains of the type 1 human immunodeficiency virus Rev protein: a dual function for an arginine-rich binding motif. *Proc. Natl. Acad. Sci. USA.* **88:** 7734-7738.